aws

Guide du développeur

Amazon Kendra

Amazon Kendra: Guide du développeur

Table of Contents

..... xiii
Qu'est-ce que c'est Amazon Kendra ? 1
Interrogation Amazon Kendra 1
Avantages de Amazon Kendra 2
Amazon Kendra Éditions 2
Tarification pour Amazon Kendra 3
Utilisez-vous pour la première fois Amazon Kendra ? 4
Comment Amazon Kendra fonctionne 5
Index dans Amazon Kendra 6
Types d'index dans Amazon Kendra 6
Ajouter des documents à un index dans Amazon Kendra 10
Utilisation de champs de document Amazon Kendra réservés ou communs 11
Récupération de réponses à partir d'index dans Amazon Kendra 13
Documents 14
Types ou formats de documents 14
Attributs ou champs du document 17
Sources de données 20
Requêtes 22
Balises 22
Balisage de ressources 23
Restrictions liées aux étiquettes 23
Configuration d'Amazon Kendra 25
Inscrivez-vous pour AWS 25
Régions et points de terminaison 26
Configuration du AWS CLI 26
Configuration du AWS SDKs 27
IAM rôles d'accès pour Amazon Kendra 28
IAM rôles pour les index 28
IAM rôles pour l' BatchPutDocumentAPI 32
IAM rôles pour les sources de données 35
Rôle de cloud privé virtuel (VPC) IAM 133
IAM rôles pour les questions fréquemment posées (FAQs) 135
IAM rôles pour les suggestions de requêtes 137
IAM rôles pour le mappage principal des utilisateurs et des groupes 138

IAM rôles pour AWS IAM Identity Center 141
IAM rôles liés aux Amazon Kendra expériences 142
IAM rôles pour l'enrichissement de documents personnalisés 145
Déploiement Amazon Kendra 150
Présentation 151
Prérequis 151
Configuration de l'exemple 152
Page de recherche principale 153
Composant de recherche 153
Composante des résultats 153
Composant Facettes 154
Composant de pagination 154
Déploiement d'une application de recherche sans code 154
Fonctionnement du moteur de recherche Experience Builder 154
Concevez et optimisez votre expérience de recherche 155
Fournir un accès à votre page de recherche 157
Configuration d'une expérience de recherche 158
Ajustement de la capacité 163
Capacité de visionnage 164
Ajouter et supprimer de la capacité 165
Amazon Kendra Capacité de classement intelligente 165
Capacité de suggestions de requêtes 166
Amazon Kendra capacité d'expérience 166
Capacité d'expérience de recherche 166
rafale de requêtes adaptative 166
Premiers pas 168
Prérequis 168
Inscrivez-vous pour un Compte AWS 168
Création d’un utilisateur doté d’un accès administratif 169
Amazon Kendra ressources : SDK AWS CLI, console 171
Commencer à utiliser la Amazon Kendra console 177
Démarrer (AWS CLI) 178
Mise en route (SDK pour Python (Boto3)) 180
Mise en route (SDK for Java) 183
Commencer à utiliser S3 (console) 187
Débuter avec MySQL (console) 188

Commencer à utiliser une source d'identité IAM Identity Center (console) 191
Modification de la source d'identité de votre IAM Identity Center 194
Création d'un index 196
Ajouter des documents directement à un index avec téléchargement par lots 201
Ajouter des documents à l'aide de l' BatchPutDocumentAPI 202
Ajouter des documents à partir d'un compartiment S3 204
Ajouter des questions fréquemment posées (FAQs) à un index 207
Création de champs d'index pour un fichier FAQ 208
Fichier CSV de base 209
Fichier CSV personnalisé 209
Fichier JSON 211
Utilisation de votre fichier FAQ 214
Fichiers de FAQ dans des langues autres que l'anglais 215
Création de champs de document personnalisés 216
Mise à jour des champs de document personnalisés 217
Contrôle de l'accès des utilisateurs aux documents à l'aide de jetons 220
Utilisation d'OpenID 221
Utilisation d'un jeton Web JSON (JWT) avec un secret partagé 224
Utilisation d'un jeton Web JSON (JWT) avec une clé publique 227
Utilisation de JSON 231
Création d'un connecteur de source de données 234
Définition d'un calendrier de mise à jour 235
Configuration d'une langue 235
Connecteurs de source de données 236
Schémas de modèles de sources de données 237
Adobe Experience Manager 632
Alfresco 643
Aurora (MySQL) 653
Aurora (PostgreSQL) 662
Amazon FSx (Fenêtres) 671
Amazon FSx (NetApp ONTAP) 680
Amazon RDS/Aurora 690
Amazon RDS (Microsoft SQL Server) 698
Amazon RDS (MySQL) 708
Amazon RDS (Oracle) 718
Amazon RDS (PostgreSQL) 727

Amazon S3 736
Amazon Kendra Explorateur Web 754
Box (Cube) 777
Confluence 786
Connecteur de source de données personnalisé 808
Dropbox 817
Drupal 827
GitHub 838
Gmail 850
Google Drive 860
IBM DB2 880
Jira 889
Microsoft Exchange 897
Microsoft OneDrive 907
Microsoft SharePoint 924
Microsoft SQL Server 960
Microsoft Teams 969
Microsoft Yammer 981
MySQL 989
Oracle Database 998
PostgreSQL 1007
Quip 1016
Salesforce 1023
ServiceNow 1043
Slack 1064
Zendesk 1075
Cartographie des champs de source de données 1085
Utilisation de champs de document Amazon Kendra réservés ou communs 17
Ajouter des documents dans des langues autres que l'anglais 1090
Configuration Amazon Kendra pour utiliser un Amazon VPC 1093
Configuration Amazon VPC 1094
Connexion à Amazon VPC 1096
Connexion à une base de données 1098
Résolution des problèmes de connexion VPC 1100
Supprimer un index, une source de données ou des documents téléchargés par lots 1103
Supprimer un index 1103

Suppression d'une source de données 1104
Suppression de documents téléchargés par lots 1107
Enrichir vos documents lors de l'ingestion 1108
Comment fonctionne l'enrichissement personnalisé des documents 1109
Opérations de base pour modifier les métadonnées 1109
Fonctions Lambda : extraire et modifier les métadonnées ou le contenu 1118
Contrats de données pour les fonctions Lambda 1127
Format de document structuré 1129
Exemple de fonction Lambda qui respecte les contrats de données 1129
Recherche dans un index 1133
Interrogation d'un index 1134
Prérequis 1135
Recherche dans un index (console) 1135
Recherche dans un index (SDK) 1136
Recherche dans un index (Postman) 1138
Recherche à l'aide d'une syntaxe de requête avancée 1140
Recherche dans les langues 1145
Récupération de passages 1149
Parcourir un index 1152
Avec les résultats de recherche 1156
Recherche tabulaire pour le HTML 1159
suggestions de requêtes 1164
Suggestions de requêtes utilisant l'historique des requêtes 1165
Suggestions de requêtes à l'aide de champs 1172
Empêcher les suggestions de certaines requêtes ou de certains contenus de champs de documents 1177
Correcteur orthographique des requêtes 1182
Utilisation du correcteur orthographique des requêtes avec des limites par défaut 1183
Filtrage et recherche par facettes 1184
Facettes 1185
Utilisation des attributs du document pour filtrer les résultats de recherche 1189
Filtrer les attributs de chaque document dans les résultats de recherche 1191
Filtrage en fonction du contexte utilisateur 1191
Filtrage par jeton utilisateur 1192
Filtrage par nom d'utilisateur et par groupe 1193
Filtrage par attribut utilisateur 1194

Filtrage du contexte utilisateur pour les documents ajoutés directement à un index 1196
Filtrage du contexte utilisateur pour les questions fréquemment posées 1197
Filtrage du contexte utilisateur pour les sources de données 1197
Réponses aux requêtes et types de réponses 1216
Réponses aux requêtes 1217
Types de réponses 1220
Réglage et tri des réponses 1225
Réglage des réponses 1225
Tri des réponses 1226
Réduction/extension des résultats de requête 1229
Réduction des résultats 1231
Choix d'un document principal par ordre de tri 1231
Document manquant : stratégie clé 1232
Élargir les résultats 1232
Interactions avec d'autres Amazon Kendra fonctionnalités 1232
Optimisation de la pertinence des recherches 1234
Réglage de la pertinence au niveau de l'indice 1236
Réglage de la pertinence au niveau de la requête 1236
Obtenir des informations grâce à l'analyse des recherches 1238
Métriques pour la recherche 1238
Taux de clics 1239
Taux de clics nul 1240
Taux de résultats de recherche nul 1240
Taux de réponse instantané 1240
Requêtes les plus fréquentes 1240
Requêtes les plus fréquentes sans aucun clic 1241
Requêtes les plus fréquentes sans aucun résultat 1241
Le haut de la page a cliqué sur les documents 1242
Nombre total de requêtes 1242
Total des documents 1242
Exemple de récupération de données métriques 1243
Des indicateurs aux informations exploitables 1244
Visualisation et génération de rapports sur les analyses de recherche 1245
Graphique du nombre total de requêtes 1245
Graphique du taux de clics 1246
Graphique du taux de clic nul 1246

Graphique du taux de résultats de recherche nuls 1246
Graphique du taux de réponse instantané 1247
Soumission de commentaires pour un apprentissage progressif 1248
Utiliser la Amazon Kendra JavaScript bibliothèque pour envoyer des commentaires 1250
Étape 1 : insérez une balise de script dans votre application Amazon Kendra de recherche 1250
Étape 2 : ajouter le jeton de commentaires aux résultats de recherche 1253
Étape 3 : tester le script de feedback 1253
Utiliser l' Amazon Kendra API pour envoyer des commentaires 1254
Ajouter des synonymes personnalisés à un index 1257
Création d'un fichier de thésaurus 1259
Ajouter un thésaurus à un index 1262
Mettre à jour un thésaurus 1266
Supprimer un thésaurus 1270
Points saillants dans les résultats de recherche 1272
Tutoriel : Création d'une solution de recherche intelligente 1273
Prérequis 1274
Étape 1 : Ajouter des documents 1275
Téléchargement de l'exemple de jeu de données 1276
Création d'un compartiment Amazon S3 1278
Création de dossiers de données et de métadonnées dans votre compartiment S3 1280
Téléchargement des données d'entrée 1283
Étape 2 : Détection des entités 1285
Exécution d'une tâche d'analyse d'entités Amazon Comprehend 1286
Étape 3 : Formatage des métadonnées 1295
Téléchargement et extraction de la sortie Amazon Comprehend 1295
Téléchargement de la sortie dans le compartiment S3 1299
Conversion de la sortie au format de métadonnées Amazon Kendra 1301
Nettoyage de votre compartiment Amazon S3 1306
Étape 4 : Création d'un index et ingestion des métadonnées 1308
Création d'un index Amazon Kendra 1308
Mise à jour du rôle IAM pour l'accès à Amazon S3 1316
Création de champs d'index de recherche personnalisés pour Amazon Kendra 1320
Ajout du compartiment Amazon S3 en tant que source de données pour l'index 1326
Synchronisation de l'index Amazon Kendra 1330
Étape 5 : Interrogation de l'index 1333

Interrogation de votre index Amazon Kendra 1333
Filtrer les résultats de recherche 1339
Étape 6 : Nettoyage 1344
Nettoyage de vos fichiers 1344
..... 1345
Surveillance et journalisation 1346
Indices de surveillance 1346
Surveillance des appels d'API Amazon Kendra avec CloudTrail 1350
Informations sur Amazon Kendra dans CloudTrail 1350
Exemple : entrées du fichier journal Amazon Kendra 1351
Surveillance des appels de l'API Amazon Kendra Intelligent Ranking avec CloudTrail 1352
Informations de classement intelligent d'Amazon Kendra dans CloudTrail 1353
Exemple : entrées du fichier journal Amazon Kendra Intelligent Ranking 1354
Surveiller Amazon Kendra avec CloudWatch 1355
Consulter les statistiques d'Amazon Kendra 1356
Création d'une alarme 1356
CloudWatch Mesures pour les tâches de synchronisation d'index 1357
Mesures pour les sources de données Amazon Kendra 1359
Mesures pour les documents indexés 1361
Surveillance d'Amazon Kendra à l'aide de journaux CloudWatch 1362
Flux de journaux des sources de données 1363
Flux de journaux de documents 1364
Afficher les statistiques Amazon Kendra pour vos tâches de synchronisation 1365
Sécurité 1368
Protection des données 1369
Chiffrement au repos 1370
Chiffrement en transit 1370
Gestion des clés 1370
Points de terminaison d’un VPC (AWS PrivateLink) 1371
Considérations relatives aux points de terminaison VPC Amazon Kendra et Amazon Kendra Intelligent Ranking 1371
Création d'un point de terminaison VPC d'interface pour Amazon Kendra et Amazon Kendra Intelligent Ranking 1371
Création d'une politique de point de terminaison VPC pour Amazon Kendra et Amazon Kendra Intelligent Ranking 1372
Gestion des identités et des accès 1374

Public ciblé 1374
Authentification par des identités 1375
Gestion des accès à l’aide de politiques 1378
Comment Amazon Kendra travaille avec IAM 1381
Exemples de politiques basées sur l’identité 1387
AWS politiques gérées 1393
Résolution des problèmes 1398
Bonnes pratiques de sécurité 1401
Application du principe du moindre privilège 1401
Autorisations de contrôle d'accès basé sur les rôles (RBAC) 1401
Journalisation et surveillance dans Amazon Kendra 1402
Validation de conformité 1402
Résilience 1403
Sécurité de l’infrastructure 1404
Analyse de la configuration et des vulnérabilités 1404
Quotas 1406
Régions prises en charge 1406
Quotas 1406
Quotas d'indice 1406
Quotas de connecteurs de sources de données 1407
FAQ sur les quotas 1409
Quotas du thésaurus 1410
Amazon Kendra quotas d'expérience 1410
Quotas de requêtes et de résultats de recherche 1411
Quotas de suggestions de requêtes 1413
Quotas de documents 1414
Quotas de résultats de recherche en vedette 1415
Rescore/ Quotas de résultats de recherche 1416
Résolution des problèmes 1418
Dépannage des sources de données 1418
Mes documents n'ont pas été indexés 1418
Ma tâche de synchronisation a échoué 1419
Ma tâche de synchronisation est incomplète 1420
Ma tâche de synchronisation a réussi mais aucun document n'est indexé 1420
Je rencontre des problèmes de format de fichier lors de la synchronisation de ma source de données 1421

Je souhaite générer un rapport d'historique de synchronisation pour mes documents 1421
Combien de temps prend la synchronisation d'une source de données ? 1423
Quels sont les frais de synchronisation d'une source de données ? 1423
Je reçois une erreur Amazon EC2 d'autorisation 1423
Je ne parviens pas à utiliser les liens de l'index de recherche pour ouvrir mes Amazon S3 objets 1424
Je reçois un message d'erreur AccessDenied lors de l'utilisation d'un fichier de certificat SSL 1424
Je reçois une erreur d'autorisation lors de l'utilisation d'une source SharePoint de données 1424
Mon index n'explore pas les documents de ma source de données Confluence 1425
Résolution des problèmes liés aux résultats de recherche documentaire 1425
Les résultats de ma recherche ne correspondent pas à ma requête 1425
Pourquoi est-ce que je ne vois que 100 résultats ? 1426
Pourquoi les documents que je m'attends à voir disparaissent-ils ? 1426
Pourquoi est-ce que je vois des documents dotés d'une politique ACL ? 1426
Dépannage de problèmes généraux 1426
Amazon Kendra Classement intelligent 1428
Classement intelligent pour l'autogestion OpenSearch 1428
Comment fonctionne le plugin de recherche intelligent 1428
Configuration du plugin de recherche intelligent 1429
Interaction avec le plugin de recherche intelligent 1435
Comparaison des OpenSearch résultats avec les Amazon Kendra résultats 1441
Classement sémantique des résultats d'un service de recherche 1442
Référence d’API 1452
Historique de la documentation 1453
..... mcdlxxi

Qu'est-ce que c'est Amazon Kendra ?

Amazon Kendra est un service géré de recherche intelligente et de recherche d'informations qui utilise le traitement du langage naturel et un modèle avancé d'apprentissage en profondeur. Contrairement à la recherche traditionnelle basée sur des mots clés, elle Amazon Kendra utilise la similarité sémantique et contextuelle, ainsi que les capacités de classement, pour déterminer si un fragment de texte ou un document est pertinent pour une requête de récupération.

Vous pouvez ainsi créer une expérience de recherche et de récupération unifiée en connectant plusieurs référentiels de données à un index et en ingérant et en analysant des documents. Amazon Kendra Vous pouvez utiliser les métadonnées de vos documents pour créer une expérience de recherche personnalisée et riche en fonctionnalités pour vos utilisateurs, afin de les aider à trouver efficacement les bonnes réponses à leurs questions.

Amazon Kendra propose un index GenAI extrêmement précis pour la récupération, la génération augmentée (RAG) et la recherche d'entreprise sur vos données. Vous pouvez utiliser les indices Kendra GenAI dans les bases de connaissances Amazon Q BusinessAmazon Bedrock pour créer des applications d'IA génératives à partir de vos données propriétaires.

> **Note**
>
> Vous pouvez également utiliser les fonctionnalités Amazon Kendra de recherche sémantique pour reclasser les résultats d'un autre service de recherche. Consultez Amazon Kendra Intelligent Ranking pour plus de détails.

Qu'est-ce que c'est Amazon Kendra ?

Interrogation Amazon Kendra

Vous pouvez poser Amazon Kendra les types de requêtes suivants :

Questions factuelles —Des questions simples sur qui, quoi, quand ou où, telles que Où se trouve le centre de service le plus proche de Seattle ? Les questions factoïdes ont des réponses factuelles qui peuvent être renvoyées sous la forme d'un seul mot ou d'une seule phrase. La réponse est extraite d'une FAQ ou de vos documents indexés.

Questions descriptives — Questions dont la réponse peut être une phrase, un passage ou un document entier. Par exemple, comment connecter mon Echo Plus à mon réseau ? Ou, Comment puis-je obtenir des avantages fiscaux pour les familles à faible revenu ?

Questions relatives aux mots clés et au langage naturel : questions qui incluent un contenu conversationnel complexe dont le sens peut ne pas être clair. Par exemple, discours liminaire. Lorsqu'il Amazon Kendra rencontre un mot tel que « adresse », qui a plusieurs significations contextuelles, il déduit correctement le sens de la requête de recherche et renvoie les informations pertinentes.

Avantages de Amazon Kendra

Amazon Kendra est hautement évolutif, capable de répondre aux exigences de performance, est étroitement intégré à d'autres AWS services tels que Amazon Q Business, Amazon S3et Amazon BedrockAmazon Lex, et offre une sécurité de niveau entreprise. Avantages liés à l'utilisation d' Amazon Kendra :

Simplicité :Amazon Kendra fournit une console et une API pour gérer les documents que vous souhaitez rechercher. Vous pouvez utiliser une API de recherche simple à Amazon Kendra intégrer à vos applications clientes, telles que des sites Web ou des applications mobiles.

Connectivité :Amazon Kendra peut se connecter à des référentiels de données ou à des sources de données tiers tels que Microsoft SharePoint. Vous pouvez facilement indexer et rechercher vos documents à l'aide de votre source de données.

Précision : contrairement aux services de recherche traditionnels qui utilisent des recherches par mots clés, ils Amazon Kendra tentent de comprendre le contexte de la question et renvoient le mot, l'extrait ou le document le plus pertinent pour votre requête. Amazon Kendra utilise l'apprentissage automatique pour améliorer les résultats de recherche au fil du temps.

Sécurité :Amazon Kendra offre une expérience de recherche d'entreprise hautement sécurisée. Les résultats de votre recherche reflètent le modèle de sécurité de votre organisation et peuvent être filtrés en fonction de l'accès des utilisateurs ou des groupes aux documents. Les clients sont responsables de l'authentification et de l'autorisation d'accès des utilisateurs.

Amazon Kendra Éditions

Amazon Kendra propose trois types d'index : GenAI Enterprise Edition, Basic Enterprise Edition et Basic Developer Edition.

L'édition GenAI Enterprise fournit une précision maximale en tirant parti des dernières technologies de récupération d'informations et des modèles sémantiques. Il offre une haute disponibilité et est conçu pour les charges de travail de production. Pour une expérience optimale, nous vous recommandons d'utiliser l'indice GenAI.

L'édition Basic Enterprise fournit des fonctionnalités de recherche sémantique et offre un service de haute disponibilité adapté aux charges de travail de production. L'édition Basic Developer offre également des fonctionnalités de recherche sémantique, conçues pour créer proof-of-concept des solutions, mais elle n'est pas recommandée pour les charges de travail de production.

Pour un aperçu de ces indices, consultez la section Types d'index.

> **Note**
>
> Pour obtenir la liste des régions, des points de terminaison et des quotas de service pris en charge par Amazon Kendra, consultez la section Amazon Kendra Points de terminaison et quotas.

Tarification pour Amazon Kendra

Vous pouvez démarrer gratuitement avec l'index Amazon Kendra GenAI Enterprise Edition ou l'index Amazon Kendra Developer Edition qui permet d'utiliser jusqu'à 750 heures pendant les 30 premiers jours.

> **Note**
>
> L'utilisation du connecteur ne donne pas droit à une utilisation gratuite.

À l'expiration de votre période d'essai, tous les Amazon Kendra index provisionnés vous sont facturés, même s'ils sont vides et qu'aucune requête n'est exécutée. Après l'expiration de la période d'essai, des frais supplémentaires sont facturés pour la numérisation et la synchronisation de documents à l'aide Amazon Kendra des sources de données.

Pour une liste complète des frais et des prix, consultez la section Amazon Kendra des prix.

Utilisez-vous pour la première fois Amazon Kendra ?

Si vous utilisez pour la première fois Amazon Kendra, nous vous recommandons de lire les sections suivantes dans l'ordre :

1	2	3	4	5	6
Comment Amazon Kendra fonctionne	Premiers pas	Création d'un index	Ajouter des documents directement à un index avec téléchargement par lots	Création d'un connecteur de source de données	Recherche dans un index
Présente Amazon Kendra les composants et décrit comment les utiliser pour créer une solution de recherche.	Explique comment configurer votre compte et tester l'API Amazon Kendra de recherche.	Explique comment Amazon Kendra créer un index de recherche et ajouter des sources de données pour synchroniser vos documents.	Explique comment ajouter des documents directement à un Amazon Kendra index.	Explique comment ajouter des documents de votre référentiel de données à un Amazon Kendra index.	Explique comment utiliser l'API Amazon Kendra de recherche pour effectuer une recherche dans un index.

Comment Amazon Kendra fonctionne

Amazon Kendra fournit une fonctionnalité de génération augmentée (RAG) de recherche et de récupération pour votre application. Il indexe vos documents directement, ou à partir de votre référentiel de documents tiers, et fournit intelligemment des informations pertinentes à vos utilisateurs. Vous pouvez l'utiliser Amazon Kendra pour créer un index actualisable de documents de différents types. Pour obtenir la liste des types de documents pris en charge par Amazon Kendra, consultez la section Types de documents.

Amazon Kendra s'intègre à d'autres services. Vous pouvez connecter un index Amazon Kendra GenAI Enterprise Edition à Amazon Q BusinessAmazon Bedrock pour créer votre solution de chat RAG. Vous pouvez également activer les robots de Amazon Lex discussion avec Amazon Kendra la recherche pour fournir des réponses utiles aux questions des utilisateurs. Vous pouvez également utiliser un Amazon Simple Storage Service bucket comme source de données pour vous connecter Amazon Kendra à vos documents et les indexer.

Amazon Kendra comporte les éléments suivants :

- Un index qui contient vos documents et les rend consultables.
- Source de données qui stocke vos documents et qui Amazon Kendra se connecte à. Vous pouvez synchroniser automatiquement une source de données avec un Amazon Kendra index afin que votre index reste à jour avec votre référentiel source.
- API d'ajout de documents qui ajoute des documents directement à un index.
- Une API de récupération qui récupère des passages ou des extraits de texte pertinents à la suite d'une requête d'entrée.
- Une API de requête qui recherche un index en fonction d'une requête d'entrée.

Vous pouvez l'utiliser Amazon Kendra via la console ou l'API. Vous pouvez créer, mettre à jour et supprimer des index. La suppression d'un index supprime tous ses connecteurs de source de données et supprime définitivement toutes les informations de votre document. Amazon Kendra

Rubriques

- Index dans Amazon Kendra
- Documents
- Sources de données

- Requêtes
- Balises

Index dans Amazon Kendra

Un index contient le contenu de vos documents et est structuré de manière à rendre les documents consultables. Cette section fournit une vue d'ensemble des types d'index et des fonctionnalités d'index pris en charge.

Rubriques

- Types d'index dans Amazon Kendra
- Ajouter des documents à un index dans Amazon Kendra
- Utilisation de champs de document Amazon Kendra réservés ou communs
- Récupération de réponses à partir d'index dans Amazon Kendra

Types d'index dans Amazon Kendra

Amazon Kendra possède trois types d'index : l'index GenAI Enterprise Edition, l'index Enterprise Edition et l'index Developer Edition. Les sections suivantes décrivent les fonctionnalités de chaque indice.

Rubriques

- Index de l'édition Enterprise d'Amazon Kendra GenAI
- Index de l'édition Amazon Kendra Enterprise
- Index de l'édition Amazon Kendra Developer

Index de l'édition Enterprise d'Amazon Kendra GenAI

Un index Amazon Kendra GenAI Enterprise Edition offre la plus grande précision pour le fonctionnement de l'API Retrieve et pour les cas d'utilisation de la génération augmentée de récupération (RAG). Il repose sur les dernières technologies de récupération d'informations, telles que la recherche hybride (mot-clé et vecteur), l'intégration sémantique et les modèles de reclassement, et a été testé sur divers ensembles de données. Le fonctionnement de l'API Query offre une précision similaire pour un index Amazon Kendra GenAI Enterprise Edition par rapport aux index Amazon Kendra Developer Edition et Amazon Kendra Enterprise Edition.

Un index Amazon Kendra GenAI Enterprise Edition permet la mobilité de vos données indexées entre AWS les services d'IA générative. Grâce à cette fonctionnalité, vous pouvez réutiliser facilement vos investissements sans avoir à reconstruire les index. Vous pouvez l'utiliser dans une base de connaissances Amazon Bedrock en tant que retriever géré, et l'intégrer aux outils Amazon Bedrock tels que les agents et les flux rapides pour créer des assistants d'intelligence artificielle avancés. Vous pouvez également l'utiliser Amazon Q Businesspour un assistant numérique entièrement géré.

Un indice Amazon Kendra GenAI Enterprise Edition propose des unités de capacité plus petites et plus granulaires et un prix de départ inférieur à celui des deux autres types d'indices. Cela vous permet d'utiliser vos capacités de manière plus efficace.

> **Note**
>
> Pour une expérience et une précision optimales, nous vous recommandons de choisir un index Amazon Kendra GenAI Enterprise Edition.

Rubriques

- Fonctionnalités prises en charge
- Limites

Fonctionnalités prises en charge

Les fonctionnalités suivantes sont prises en charge pour un index Amazon Kendra GenAI Enterprise Edition si vous utilisez l'opération d'API Retrieve pour les cas d'utilisation de RAG :

- Support complet : compartiments de scores de confiance, filtrage, facettage, réglage de la pertinence, enrichissement personnalisé des documents, métadonnées personnalisées et ajustement de la capacité des requêtes et de la capacité des documents.
- Support partiel : connecteurs de source de données et filtrage du contexte utilisateur. Pour plus d'informations sur les fonctionnalités partiellement prises en charge, consultez la section Limitations.

Les fonctionnalités suivantes sont prises en charge pour un index Amazon Kendra GenAI Enterprise Edition si vous utilisez l'opération Query API pour des cas d'utilisation de recherche :

- Support complet : classement des documents, réponses extractives aux questions, compartiments de scores de confiance, filtrage, facettage, tri, réduction et extension des résultats des requêtes, navigation par index, requêtes booléennes, correspondance exacte, requêtesgénériques, suggestions de requêtes, correcteur orthographique, réglage de la pertinence, apprentissage progressif, document personnalisé https://docs.aws.amazon.com/kendra/latest/dg/custom-document-enrichment.htmlenrichissement, métadonnées personnalisées, ajustement de la capacité de requête et de la capacité des documents, et expérience de recherche.
- Support partiel : connecteurs de source de données et filtrage du contexte utilisateur. Pour plus d'informations sur les fonctionnalités partiellement prises en charge, consultez la section Limitations.

Limites

Ce qui suit décrit les limites connues d'un index Amazon Kendra GenAI Enterprise Edition :

- Les index Amazon Kendra GenAI Enterprise Edition ne sont disponibles que dans l'est des États-Unis (Virginie du Nord) et dans l'ouest des États-Unis (Oregon).
- Les index Amazon Kendra GenAI Enterprise Edition ne prennent en charge que le contenu en anglais.
- Les index Amazon Kendra GenAI Enterprise Edition ne prennent en charge que les connecteurs de source de données Amazon Kendra v2.0.
- Dans un index Amazon Kendra GenAI Enterprise Edition, vous ne pouvez utiliser les attributs utilisateur que pour filtrer les résultats de recherche en fonction du contexte utilisateur.
- Les index Amazon Kendra GenAI Enterprise Edition ne prennent pas en charge le contrôle d'accès utilisateur basé sur des jetons ni le contrôle d'accès utilisateur aux documents basé sur l'ID utilisateur et le groupe.
- Le fonctionnement de l'CreateAccessControlConfigurationAPI est désactivé pour les index Amazon Kendra GenAI Enterprise Edition.
- Si vous utilisez un index Amazon Kendra GenAI Enterprise Edition avec Amazon Q Business, notez ce qui suit concernant le contrôle de l'accès des utilisateurs finaux aux documents :

  Amazon Q Business utilise l'adresse e-mail de l'utilisateur pour déterminer l'accès de l'utilisateur final aux documents d'un index. Lorsque vous connectez un index Amazon Kendra à Amazon Kendra Amazon Q Business, il Amazon Q Business transmet l'adresse e-mail d'identification de l'utilisateur à Amazon Kendra afin de permettre le filtrage des documents pour les utilisateurs finaux. Si les sources de données connectées à votre index Amazon Kendra n'utilisent pas le

filtrage des documents basé sur l'identifiant d'e-mail, ou si l'identifiant d'e-mail n'est pas présent, Amazon Q Business génère des réponses uniquement à partir de documents publics.

Index de l'édition Amazon Kendra Enterprise

Un index Amazon Kendra Enterprise Edition fournit des fonctionnalités de recherche sémantique et propose un service de haute disponibilité adapté aux charges de travail de production.

Rubriques

- Fonctionnalités prises en charge
- Limites

Fonctionnalités prises en charge

Les fonctionnalités suivantes sont prises en charge pour un index Amazon Kendra Enterprise Edition si vous utilisez l'opération d'API Retrieve pour les cas d'utilisation de RAG : requêtes utilisant une syntaxe de requête avancée, corrections orthographiques suggérées pour les requêtes, facettage, suggestions de requêtes pour compléter automatiquement les requêtes de recherche et apprentissage progressif.

Toutes les fonctionnalités sont prises en charge pour un index Amazon Kendra Enterprise Edition si vous utilisez l'opération Query API pour les cas d'utilisation de la recherche.

Limites

Ce qui suit décrit les limites connues d'un index Amazon Kendra Enterprise Edition :

- Si vous utilisez un index Amazon Kendra Enterprise Edition avec Amazon Q Business, notez ce qui suit concernant le contrôle de l'accès des utilisateurs finaux aux documents :

  Amazon Q Business utilise l'adresse e-mail de l'utilisateur pour déterminer l'accès de l'utilisateur final aux documents d'un index. Lorsque vous connectez un index Amazon Kendra à Amazon Kendra Amazon Q Business, il Amazon Q Business transmet l'adresse e-mail d'identification de l'utilisateur à Amazon Kendra afin de permettre le filtrage des documents pour les utilisateurs finaux. Si les sources de données connectées à votre index Amazon Kendra n'utilisent pas le filtrage des documents basé sur l'identifiant d'e-mail, ou si l'identifiant d'e-mail n'est pas présent, Amazon Q Business génère des réponses uniquement à partir de documents publics.

Index de l'édition Amazon Kendra Developer

Un index Amazon Kendra Developer Edition fournit également des fonctionnalités de recherche sémantique vous permettant de tester vos cas d'utilisation. Cependant, nous ne le recommandons pas pour les cas d'utilisation en production.

Rubriques

- Fonctionnalités prises en charge
- Limites

Fonctionnalités prises en charge

Les fonctionnalités suivantes sont prises en charge pour un index Amazon Kendra Developer Edition si vous utilisez l'opération de l'API Retrieve pour les cas d'utilisation de RAG : requêtes utilisant une syntaxe de requête avancée, corrections orthographiques suggérées pour les requêtes, facettage, suggestions de requêtes pour compléter automatiquement les requêtes de recherche et apprentissage progressif.

Toutes les fonctionnalités sont prises en charge pour un index Amazon Kendra Developer Edition si vous utilisez l'opération Query API pour les cas d'utilisation de la recherche.

Limites

Ce qui suit décrit les limites connues d'un index Amazon Kendra Developer Edition :

- Si vous utilisez un index Amazon Kendra Developer Edition avec Amazon Q Business, notez ce qui suit concernant le contrôle de l'accès des utilisateurs finaux aux documents :

  Amazon Q Business utilise l'adresse e-mail de l'utilisateur pour déterminer l'accès de l'utilisateur final aux documents d'un index. Lorsque vous connectez un index Amazon Kendra à Amazon Kendra Amazon Q Business, il Amazon Q Business transmet l'adresse e-mail d'identification de l'utilisateur à Amazon Kendra afin de permettre le filtrage des documents pour les utilisateurs finaux. Si les sources de données connectées à votre index Amazon Kendra n'utilisent pas le filtrage des documents basé sur l'identifiant d'e-mail, ou si l'identifiant d'e-mail n'est pas présent, Amazon Q Business génère des réponses uniquement à partir de documents publics.

Ajouter des documents à un index dans Amazon Kendra

La façon dont vous ajoutez des documents à un index dépend de la manière dont vous les stockez.

- Si vous stockez vos documents dans un référentiel, tel qu'un Amazon S3 bucket ou un SharePoint site Microsoft, vous utilisez un connecteur de source de données pour indexer vos documents à partir de votre référentiel.
- Si vous ne stockez pas vos documents dans un référentiel, vous utilisez l'opération BatchPutDocumentAPI pour les indexer directement.
- Pour les questions et réponses aux FAQ, qui doivent être stockées dans un compartiment Amazon Kendra (Amazon S3), vous les téléchargez depuis le compartiment.

Vous pouvez créer des index à l'aide de la Amazon Kendra console AWS CLI, du ou d'un AWS SDK. Pour plus d'informations sur les types de documents pouvant être indexés, consultez la section Types de documents.

Utilisation de champs de document Amazon Kendra réservés ou communs

Grâce au fonctionnement de l'UpdateIndexAPI, vous pouvez créer des champs réservés ou communs. Pour ce faire, vous devez utiliser `DocumentMetadataConfigurationUpdates` et spécifier le nom du champ d'index Amazon Kendra réservé à mapper à votre attribut/nom de champ de document équivalent. Vous pouvez également créer des champs personnalisés.

Si vous utilisez un connecteur de source de données, la plupart incluent des mappages de champs qui font correspondre les champs de votre document de source de données aux champs d' Amazon Kendra index. Si vous utilisez la console, vous mettez à jour les champs en sélectionnant votre source de données, en sélectionnant l'action de modification, puis en passant à côté de la section des mappages de champs pour configurer la source de données.

Vous pouvez configurer l'`Search`objet pour définir un champ comme affichable, facetable, consultable ou triable. Vous pouvez configurer l'`Relevance`objet pour définir l'ordre de classement d'un champ, la durée d'augmentation ou la période à appliquer aux valeurs de renforcement, de fraîcheur, de valeur d'importance et d'importance mappées à des valeurs de champ spécifiques.

Si vous utilisez la console, vous pouvez configurer les paramètres de recherche d'un champ en sélectionnant l'option à facettes dans le menu de navigation. Pour définir le réglage de la pertinence, sélectionnez l'option permettant de rechercher votre index dans le menu de navigation, entrez une requête et utilisez les options du panneau latéral pour ajuster la pertinence de la recherche. Vous ne pouvez pas modifier le type de champ une fois que vous l'avez créé.

Amazon Kendrapossède les champs de document réservés ou communs suivants que vous pouvez utiliser :

- `_authors`— Une liste d'un ou de plusieurs auteurs responsables du contenu du document.
- `_category`— Catégorie qui place un document dans un groupe spécifique.
- `_created_at`— Date et heure au format ISO 8601 auxquelles le document a été créé. Par exemple, 2012-03-25T 12:30:10 + 01:00 est le format date-heure ISO 8601 pour le 25 mars 2012 à 12h30 (plus 10 secondes) en heure d'Europe centrale.
- `_data_source_id`— Identifiant de la source de données contenant le document.
- `_document_body`— Le contenu du document.
- `_document_id`— Identifiant unique du document.
- `_document_title`— Le titre du document.
- `_excerpt_page_number`— Le numéro de page d'un fichier PDF où apparaît l'extrait du document. Si votre index a été créé avant le 8 septembre 2020, vous devez réindexer vos documents avant de pouvoir utiliser cet attribut.
- `_faq_id`— S'il s'agit d'un document de type question-réponse (FAQ), un identifiant unique pour la FAQ.
- `_file_type`— Le type de fichier du document, tel que pdf ou doc.
- `_last_updated_at`— Date et heure au format ISO 8601 auxquelles le document a été mis à jour pour la dernière fois. Par exemple, 2012-03-25T 12:30:10 + 01:00 est le format date-heure ISO 8601 pour le 25 mars 2012 à 12h30 (plus 10 secondes) en heure d'Europe centrale.
- `_source_uri`— L'URI où le document est disponible, par exemple, l'URI du document sur le site Web d'une entreprise.
- `_version`— Identifiant pour la version spécifique d'un document.
- `_view_count`— Le nombre de fois que le document a été consulté.
- `_language_code`(String) — Code d'une langue qui s'applique au document. La valeur par défaut est l'anglais si vous ne spécifiez aucune langue. Pour plus d'informations sur les langues prises en charge, y compris leurs codes, voir Ajout de documents dans des langues autres que l'anglais.

Vous créez des champs personnalisés à l'`DocumentMetadataConfigurationUpdates`aide de l'opération `UpdateIndex` API, comme vous le faites lors de la création d'un champ réservé ou commun. Vous devez définir le type de données approprié pour votre champ personnalisé.

Si vous utilisez la console, vous mettez à jour les champs en sélectionnant votre source de données, en sélectionnant l'action de modification, puis en passant à côté de la section des mappages de champs pour configurer la source de données. Certaines sources de données ne prennent pas en

charge l'ajout de nouveaux champs ou de champs personnalisés. Vous ne pouvez pas modifier le type de champ une fois que vous l'avez créé.

Les types que vous pouvez définir pour les champs personnalisés sont les suivants :

- Date
- Nombre
- Chaîne
- Liste de chaînes

Si vous avez ajouté des documents à un index à l'aide de l'opération BatchPutDocumentAPI, `Attributes` répertorie les champs/attributs de vos documents et vous créez des champs à l'aide de l'`DocumentAttribute`objet.

Pour les documents indexés à partir d'une source de Amazon S3 données, vous créez des champs à l'aide d'un fichier de métadonnées JSON qui inclut les informations des champs.

Si vous utilisez une base de données prise en charge comme source de données, vous pouvez configurer vos champs à l'aide de l'option de mappage de champs.

Récupération de réponses à partir d'index dans Amazon Kendra

Après avoir créé un index, vous pouvez commencer à rechercher dans vos documents.

Pour rechercher un Amazon Kendra index, vous devez utiliser l'opération d'API Retrieve ou l'opération d'API Query.

L'opération de l'API Retrieve est idéale pour les cas d'utilisation de la génération augmentée de récupération (RAG). Pour une requête donnée, il renvoie une liste classée de passages sémantiquement pertinents contenant jusqu'à 200 mots symboliques. Vous pouvez les envoyer à un grand modèle de langage (LLM) pour générer une réponse à l'aide de RAG. Pour plus d'informations, consultez la section Recherche dans un index.

L'opération Query API est idéale pour les cas d'utilisation de la recherche de documents. Pour une requête donnée, il renvoie une liste de documents classés avec des extraits de 100 mots pertinents pour la requête. Cela est utile pour les cas d'utilisation traditionnels de la recherche de documents où les utilisateurs parcourent une liste de documents classés.

Pour connaître les fonctionnalités prises en charge par les opérations des API Retrieve et Query pour chaque type d'index, consultez la section Types d'index.

Documents

Cette section explique comment Amazon Kendra indexer les nombreux formats de documents qu'il prend en charge et les différents champs/attributs des documents.

Rubriques

- Types ou formats de documents
- Attributs ou champs du document

Types ou formats de documents

Amazon Kendra prend en charge les types ou formats de documents courants tels que PDF, HTML PowerPoint, Word, etc. Un index peut contenir plusieurs formats de document.

Amazon Kendra extrait le contenu des documents afin de les rendre consultables. Les documents sont analysés de manière à optimiser la recherche sur le texte extrait et sur tout contenu tabulaire (tableaux HTML) dans les documents. Cela implique de structurer les documents en champs ou attributs utilisés pour la recherche. Les métadonnées du document, telles que la date de dernière modification, peuvent être des champs utiles pour la recherche.

Les documents peuvent être organisés en lignes et en colonnes. Par exemple, chaque document est une ligne et chaque champ/attribut du document, tel que le titre et le corps du contenu, est une colonne. Par exemple, si vous utilisez une base de données comme source de données, les données doivent être structurées ou organisées en lignes et en colonnes.

Vous pouvez ajouter des documents à votre index de différentes manières :

- API BatchPutDocument
- Connecteur de source de données

Si vous souhaitez ajouter un fichier de FAQ, vous devez utiliser l'CreateFaqAPI pour ajouter le fichier stocké dans un Amazon S3 bucket. Vous pouvez choisir entre un format CSV de base, un format CSV qui inclut des champs/attributs personnalisés dans un en-tête et un format JSON qui inclut des champs personnalisés. Le format par défaut est le CSV de base.

Vous trouverez ci-dessous des informations sur chaque format de document pris en charge et sur le traitement Amazon Kendra de chaque format lors de l'indexation de documents.

Format du document	Traité comme	Comment le document est traité	Structure originale
Format de document portable (PDF)	HTML	Converti en HTML, puis le contenu est extrait.	Non structuré
HyperText Langage de balisage (HTML)	HTML	Les balises HTML sont filtrées pour extraire le contenu. Le contenu doit se situer entre les balises de `HTML` début et de fermeture principales (`<HTML>content</HTML>`).	Semi-structuré
XML (Extensible Markup Language)	xml	Les balises XML sont filtrées pour extraire le contenu.	Semi-structuré
Transformation du langage des feuilles de style extensibles (XSLT)	XSLT	Les balises sont filtrées pour extraire le contenu.	Semi-structuré
MarkDown (MARYLAND)	Texte brut	Le contenu est extrait avec MarkDown la syntaxe incluse.	Semi-structuré
Valeurs séparées par des virgules (CSV)	CSV	Contenu extrait de chaque cellule, avec un seul fichier traité comme un résultat de document unique.	Structuré pour les fichiers FAQ, sinon semi-structuré
Microsoft Excel (XLS et XLSX)	XLS et XLSX	Contenu extrait de chaque cellule, avec	Semi-structuré

Format du document	Traité comme	Comment le document est traité	Structure originale
		un seul fichier traité comme un résultat de document unique.	
JavaScript Notation d'objets (JSON)	Texte brut	Le contenu est extrait avec la syntaxe JSON incluse.	Semi-structuré
Format de texte enrichi (RTF)	RTF	La syntaxe RTF est filtrée pour extraire le contenu.	Semi-structuré
Microsoft PowerPoint (PPT)	PPT, PPTX	Seul le contenu textuel est extrait des PowerPoint diapositives à des fins de recherche. Les images et autres contenus ne sont pas extraits.	Non structuré
Microsoft Word	DOC, DOCX	Seul le contenu textuel est extrait des pages Word à des fins de recherche. Les images et autres contenus ne sont pas extraits.	Non structuré
Texte brut (TXT)	TXT	Tout le texte du document texte est extrait.	Non structuré

Attributs ou champs du document

Des attributs ou des champs sont associés à un document. Les champs d'un document sont les propriétés d'un document ou le contenu de la structure d'un document. Par exemple, chacun de vos documents peut contenir le titre, le corps du texte et l'auteur. Vous pouvez également ajouter des champs personnalisés pour vos documents spécifiques. Par exemple, si votre index recherche des documents fiscaux, vous pouvez spécifier un champ personnalisé pour le type de document fiscal tel que W-2, 1099, etc.

Avant de pouvoir utiliser un champ de document dans une requête, il doit être mappé à un champ d'index. Par exemple, le champ de titre peut être mappé au champ`_document_title`. Pour plus d'informations, consultez la section Mappage des champs. Pour ajouter un nouveau champ, vous devez créer un champ d'index auquel mapper le champ. Vous créez des champs d'index à l'aide de la console ou de l'UpdateIndexAPI.

Vous pouvez utiliser les champs du document pour filtrer les réponses et créer des résultats de recherche à facettes. Par exemple, vous pouvez filtrer une réponse pour renvoyer uniquement une version spécifique d'un document, ou vous pouvez filtrer les recherches pour ne renvoyer que les documents fiscaux de type 1099 correspondant au terme de recherche. Pour plus d'informations, consultez la section Filtrage et recherche par facettes.

Vous pouvez également utiliser les champs du document pour ajuster manuellement la réponse à la requête. Par exemple, vous pouvez choisir d'augmenter l'importance du champ de titre pour augmenter le poids qui lui est Amazon Kendra attribué lors de la détermination des documents à renvoyer dans la réponse. Pour plus d'informations, consultez la section Optimisation de la pertinence de la recherche.

Si vous ajoutez un document directement à un index, vous devez spécifier les champs dans le paramètre d'entrée Document de l'BatchPutDocumentAPI. Vous spécifiez les valeurs des champs personnalisés dans un tableau d'DocumentAttributeobjets. Si vous utilisez une source de données, la méthode que vous utilisez pour ajouter les champs du document dépend de la source de données. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.

Utilisation de champs de document Amazon Kendra réservés ou communs

Avec l'UpdateIndex API, vous pouvez créer des champs réservés ou communs en utilisant `DocumentMetadataConfigurationUpdates` et en spécifiant le nom du champ d'index Amazon Kendra réservé à mapper à l'attribut/nom de champ de votre document équivalent. Vous pouvez

également créer des champs personnalisés. Si vous utilisez un connecteur de source de données, la plupart incluent des mappages de champs qui font correspondre les champs de votre document de source de données aux champs d' Amazon Kendra index. Si vous utilisez la console, vous mettez à jour les champs en sélectionnant votre source de données, en sélectionnant l'action de modification, puis en passant à côté de la section des mappages de champs pour configurer la source de données.

Vous pouvez configurer l'`Search`objet pour définir un champ comme affichable, facetable, consultable ou triable. Vous pouvez configurer l'`Relevance`objet pour définir l'ordre de classement d'un champ, la durée d'augmentation ou la période à appliquer aux valeurs de renforcement, de fraîcheur, de valeur d'importance et d'importance mappées à des valeurs de champ spécifiques. Si vous utilisez la console, vous pouvez définir les paramètres de recherche d'un champ en sélectionnant l'option à facettes dans le menu de navigation. Pour définir le réglage de la pertinence, sélectionnez l'option permettant de rechercher votre index dans le menu de navigation, entrez une requête et utilisez les options du panneau latéral pour ajuster la pertinence de la recherche. Vous ne pouvez pas modifier le type de champ une fois que vous l'avez créé.

Amazon Kendra possède les champs de document réservés ou communs suivants que vous pouvez utiliser :

- `_authors`—Une liste d'un ou de plusieurs auteurs responsables du contenu du document.
- `_category`: catégorie qui place un document dans un groupe spécifique.
- `_created_at`: date et heure au format ISO 8601 auxquelles le document a été créé. Par exemple, 2012-03-25T12:30:10+01:00 est le format de date et d'heure ISO 8601 pour le 25 mars 2012 à 12 h 30 (plus 10 secondes) à l'heure d'Europe centrale.
- `_data_source_id`: l'identifiant de la source de données qui contient le document.
- `_document_body`: le contenu du document.
- `_document_id`—Un identifiant unique pour le document.
- `_document_title`: le titre du document.
- `_excerpt_page_number`: le numéro de page d'un fichier PDF où apparaît l'extrait du document. Si votre index a été créé avant le 8 septembre 2020, vous devez réindexer vos documents avant de pouvoir utiliser cet attribut.
- `_faq_id`—S'il s'agit d'un document de type question-réponse (FAQ), un identifiant unique pour la FAQ.
- `_file_type`: le type de fichier du document, tel que pdf ou doc.

- `_last_updated_at`: date et heure au format ISO 8601 auxquelles le document a été mis à jour pour la dernière fois. Par exemple, 2012-03-25T12:30:10+01:00 est le format de date et d'heure ISO 8601 pour le 25 mars 2012 à 12 h 30 (plus 10 secondes) à l'heure d'Europe centrale.
- `_source_uri`: l'URI où le document est disponible. Par exemple, l'URI du document sur le site Web d'une entreprise.
- `_version`—Identifiant pour la version spécifique d'un document.
- `_view_count`: le nombre de fois que le document a été consulté.
- `_language_code`(String) : code d'une langue qui s'applique au document. La valeur par défaut est l'anglais si vous ne spécifiez aucune langue. Pour plus d'informations sur les langues prises en charge, y compris leurs codes, voir Ajout de documents dans des langues autres que l'anglais.

Pour les champs personnalisés, vous pouvez les créer à l'`DocumentMetadataConfigurationUpdates`aide de l'`UpdateIndex`API, comme vous le faites lorsque vous créez un champ réservé ou commun. Vous devez définir le type de données approprié pour votre champ personnalisé. Si vous utilisez la console, vous mettez à jour les champs en sélectionnant votre source de données, en sélectionnant l'action de modification, puis en passant à côté de la section des mappages de champs pour configurer la source de données. Certaines sources de données ne prennent pas en charge l'ajout de nouveaux champs ou de champs personnalisés. Vous ne pouvez pas modifier le type de champ une fois que vous l'avez créé.

Les types que vous pouvez définir pour les champs personnalisés sont les suivants :

- Date
- Nombre
- Chaîne
- Liste de chaînes

Si vous avez ajouté des documents à l'index à l'aide de l'BatchPutDocumentAPI, `Attributes` répertorie les champs/attributs de vos documents et vous créez des champs à l'aide de l'`DocumentAttribute`objet.

Pour les documents indexés à partir d'une source de Amazon S3 données, vous créez des champs à l'aide d'un fichier de métadonnées JSON qui inclut les informations des champs.

Si vous utilisez une base de données prise en charge comme source de données, vous pouvez configurer vos champs à l'aide de l'option de mappage de champs.

Sources de données

Une source de données est un référentiel ou un emplacement de données qui Amazon Kendra se connecte à vos documents ou à votre contenu et les indexe. Par exemple, vous pouvez configurer pour vous connecter Amazon Kendra à Microsoft SharePoint afin d'explorer et d'indexer vos documents stockés dans cette source. Vous pouvez également indexer des pages Web en fournissant le URLs formulaire Amazon Kendra à explorer. Vous pouvez synchroniser automatiquement une source de données avec un Amazon Kendra index afin que les documents ajoutés, mis à jour ou supprimés dans la source de données soient également ajoutés, mis à jour ou supprimés dans l'index.

Les sources de données prises en charge sont les suivantes :

- Gestionnaire d'expérience Adobe
- En plein air
- Aurora (MySQL)
- Aurora (PostgreSQL)
- Amazon FSx (Fenêtres)
- Amazon FSx (NetApp ONTAP)
- Sources de données de base de données
- Amazon RDS (Microsoft SQL Server)
- Amazon RDS (MySQL)
- Amazon RDS (Oracle)
- Amazon RDS (PostgreSQL)
- Amazon S3 seaux
- Amazon Kendra robot d'exploration Web
- Box (Cube)
- Confluence
- Sources de données personnalisées
- Dropbox
- Drupal
- GitHub

- Gmail
- Disques durs Google Workspace
- IBM DB2
- Jira
- Microsoft Exchange
- Microsoft OneDrive
- Microsoft SharePoint
- Microsoft Teams
- Microsoft SQL Server
- Microsoft Yammer
- MySQL
- Base de données Oracle
- PostgreSQL
- Quip
- Salesforce
- ServiceNow
- Slack
- Zendesk

Pour une liste des types de documents ou des formats pris en charge par la Amazon Kendra section Types de documents. Vous devez d'abord créer un index avant de créer un connecteur de source de données pour indexer vos documents à partir de votre source de données.

> **Note**
>
> Pour créer un index de documents, il n'est pas nécessaire d'utiliser une source de données. Vous pouvez ajouter des documents directement à un index grâce au téléchargement par lots. Pour plus d'informations, consultez la section Ajout de documents directement à un index.

Pour une présentation détaillée de l'utilisation de la Amazon Kendra console, de la AWS CLI ou SDKs consultez Getting started.

Requêtes

Pour obtenir des réponses, les utilisateurs interrogent un index. Les utilisateurs peuvent utiliser le langage naturel dans leurs requêtes. La réponse contient des informations, telles que le titre, un extrait de texte et l'emplacement des documents dans l'index qui fournissent la meilleure réponse.

Amazon Kendra utilise toutes les informations que vous fournissez sur vos documents, et pas seulement le contenu des documents, pour déterminer si un document est pertinent pour la requête. Par exemple, si votre index contient des informations sur la date de dernière mise à jour des documents, vous pouvez indiquer Amazon Kendra d'attribuer une plus grande pertinence aux documents mis à jour plus récemment.

Une requête peut également contenir des critères permettant de filtrer la réponse afin que Amazon Kendra seuls les documents qui répondent aux critères de filtrage soient renvoyés. Par exemple, si vous avez créé un champ d'index appelé département, vous pouvez filtrer la réponse afin que seuls les documents dont le champ département est défini sur légal soient renvoyés. Pour plus d'informations, consultez la section Filtrage de la recherche.

Vous pouvez influencer les résultats d'une requête en ajustant la pertinence des différents champs de l'index. Le réglage modifie l'importance d'un champ sur les résultats. Par exemple, si vous soulignez l'importance des documents avec la catégorie nouveau, les documents de cette catégorie sont plus susceptibles d'être inclus dans la réponse. Pour plus d'informations, consultez la section Optimisation de la pertinence de la recherche.

Pour plus d'informations sur l'utilisation des requêtes, consultez la section Recherche dans un index.

Balises

Gérez vos index, vos sources de données et FAQs en attribuant des balises ou des étiquettes. Vous pouvez utiliser des balises pour classer vos Amazon Kendra ressources de différentes manières. Par exemple, par objectif, par propriétaire ou par application, ou par toute combinaison des deux. Chaque balise est constituée d'une clé et d'une valeur que vous définissez toutes deux.

Les balises vous aident à :

- Identifiez et organisez vos AWS ressources. De nombreux AWS services prennent en charge le balisage. Vous pouvez donc attribuer le même tag aux ressources de différents services pour indiquer que les ressources sont liées. Par exemple, vous pouvez étiqueter un index et le Amazon Lex bot qui utilise l'index avec la même balise.

- Répartir les coûts. Vous activez les tags sur le AWS Billing and Cost Management tableau de bord. AWS utilise des balises pour classer vos coûts et vous fournir un rapport mensuel de répartition des coûts. Pour plus d'informations, consultez la section Allocation et balisage des coûts dans About AWS Billing and Cost Management.
- Contrôler l'accès à vos ressources. Vous pouvez utiliser des balises dans les politiques AWS Identity and Access Management (IAM) qui contrôlent l'accès aux Amazon Kendra ressources. Vous pouvez associer ces politiques à un IAM rôle ou à un utilisateur pour activer le contrôle d'accès basé sur des balises. Pour plus d'informations, consultez la section Autorisation basée sur les balises.

Vous pouvez créer et gérer des balises à l'aide de l'API AWS Management Console, de la AWS Command Line Interface (AWS CLI) ou de l' Amazon Kendra API.

Balisage de ressources

Si vous utilisez la Amazon Kendra console, vous pouvez étiqueter les ressources lorsque vous les créez ou que vous les ajoutez ultérieurement. Vous pouvez également utiliser la console pour mettre à jour ou supprimer des balises.

Si vous utilisez le AWS Command Line Interface (AWS CLI) ou l' Amazon Kendra API, utilisez les opérations suivantes pour gérer les balises de vos ressources :

- CreateDataSource—Appliquez des balises lorsque vous créez une source de données.
- CreateFaq—Appliquez des balises lorsque vous créez une FAQ.
- CreateIndex—Appliquez des balises lorsque vous créez un index.
- ListTagsForResource—Afficher les balises associées à une ressource.
- TagResource—Ajoutez et modifiez les balises d'une ressource.
- UntagResource—Supprime les balises d'une ressource.

Restrictions liées aux étiquettes

Les restrictions suivantes s'appliquent aux balises sur les Amazon Kendra ressources :

- Nombre maximum de balises : 50
- Longueur de clé maximale : 128 caractères
- Longueur maximale de la valeur : 256 caractères

- Caractères valides pour la clé et la valeur : a—z, A—Z, espace, et les caractères suivants : _. :/= + - et @
- Les clés et les valeurs sont sensibles à la casse.
- N'utilisez pas `aws:` comme préfixe pour les clés ; seul AWS peut utiliser cette valeur.

Configuration d'Amazon Kendra

Avant d'utiliser Amazon Kendra, vous devez disposer d'un compte Amazon Web Services (AWS). Une fois que vous avez AWS créé un compte, vous pouvez accéder à Amazon Kendra via la console Amazon Kendra, AWS Command Line Interface le AWS CLI() ou le. AWS SDKs

Ce guide inclut des exemples pour AWS CLI Java et Python.

Rubriques

- Inscrivez-vous pour AWS
- Régions et points de terminaison
- Configuration du AWS CLI
- Configuration du AWS SDKs

Inscrivez-vous pour AWS

Lorsque vous vous inscrivez à Amazon Web Services (AWS), votre compte est automatiquement inscrit à tous les services AWS, y compris Amazon Kendra. Seuls les services que vous utilisez vous sont facturés.

Si vous avez déjà un AWS compte, passez à la tâche suivante. Si vous n'avez pas de compte AWS , observez la procédure suivante pour en créer un.

Pour vous inscrire à AWS

1. Ouvrez https://aws.amazon.com, puis choisissez Créer un AWS compte.
2. Suivez les instructions à l'écran pour réaliser la création du compte. Notez votre numéro de AWS compte à 12 chiffres. Dans le cadre de la procédure d'inscription, vous recevrez un appel téléphonique et vous saisirez un code PIN en utilisant le clavier numérique du téléphone.
3. Créez un utilisateur administrateur AWS Identity and Access Management (IAM). Pour obtenir des instructions, consultez Création de votre premier groupe d'utilisateurs IAM dans le guide de l'utilisateur AWS Identity and Access Management .

Régions et points de terminaison

Un point de terminaison est une URL qui est le point d'entrée d'un service Web. Chaque point final est associé à une AWS région spécifique. Si vous utilisez une combinaison de la console Amazon Kendra, du AWS CLI, et d'Amazon SDKs Kendra, faites attention à leurs régions par défaut, car tous les composants Amazon Kendra d'une campagne donnée (index, requête, etc.) doivent être créés dans la même région. Pour connaître les régions et les points de terminaison pris en charge par Amazon Kendra, consultez la section Régions et points de terminaison.

Configuration du AWS CLI

L'interface de ligne de AWS commande (AWS CLI) est un outil de développement unifié permettant de gérer AWS des services, notamment Amazon Kendra. Nous vous recommandons de l'installer.

1. Pour l'installer AWS CLI, suivez les instructions de la section Installation de l'interface de ligne de AWS commande dans le guide de l'utilisateur de l'interface de ligne de AWS commande.
2. Pour configurer le AWS CLI et configurer un profil pour l'appeler AWS CLI, suivez les instructions de la section Configuration du guide de AWS CLI l'utilisateur de l'interface de ligne de AWS commande.
3. Pour vérifier que le AWS CLI profil est correctement configuré, exécutez la commande suivante :

   ```
   aws configure --profile default
   ```

   Si votre profil a été correctement configuré, vous devez obtenir un résultat similaire à ce qui suit :

   ```
   AWS Access Key ID [****************52FQ]:
   AWS Secret Access Key [****************xgyZ]:
   Default region name [us-west-2]:
   Default output format [json]:
   ```

4. Pour vérifier que le AWS CLI est configuré pour être utilisé avec Amazon Kendra, exécutez les commandes suivantes :

   ```
   aws kendra help
   ```

   Si elle AWS CLI est correctement configurée, vous verrez une liste des AWS CLI commandes prises en charge pour Amazon Kendra, Amazon Kendra Runtime et Amazon Kendra events.

Configuration du AWS SDKs

Téléchargez et installez celui AWS SDKs que vous souhaitez utiliser. Ce guide fournit des exemples pour Python. Pour plus d'informations sur les autres AWS SDKs, consultez la section Outils pour Amazon Web Services.

Le package du SDK Python s'appelle Boto3.

Avant d'exécuter les commandes Python ci-dessous, vous devez d'abord télécharger et installer Python 3.6 ou version ultérieure pour votre système d'exploitation. Support pour Python 3.5 et versions antérieures est obsolète. Si pip n'est pas inclus dans votre répertoire de scripts Python, vous pouvez télécharger get-pip.py et le stocker dans votre répertoire de scripts. Vous pouvez également définir votre répertoire Python en tant que chemin ou variable d'environnement à l'aide d'un programme de terminal.

```
# Install the latest Boto3 release via pip
pip install boto3

# You can install a specific version of Boto3 for compatibility reasons
# Install Boto3 version 1.0 specifically
pip install boto3==1.0.0

# Make sure Boto3 is no older than version 1.15.0
pip install boto3>=1.15.0

# Avoid versions of Boto3 newer than version 1.15.3
pip install boto3<=1.15.3
```

Pour utiliser Boto3, vous devez configurer les informations d'authentification pour votre AWS compte à l'aide de la console IAM.

IAM rôles d'accès pour Amazon Kendra

Lorsque vous créez un index, une source de données ou une FAQ, Amazon Kendra vous devez accéder aux AWS ressources requises pour créer la Amazon Kendra ressource. Vous devez créer une politique AWS Identity and Access Management (IAM) avant de créer la Amazon Kendra ressource. Lorsque vous appelez l'opération, vous fournissez le nom de ressource Amazon (ARN) du rôle avec la politique jointe. Par exemple, si vous appelez l'BatchPutDocumentAPI pour ajouter des documents à partir d'un Amazon S3 compartiment, vous Amazon Kendra attribuez un rôle doté d'une politique d'accès au compartiment.

Vous pouvez créer un nouveau IAM rôle dans la Amazon Kendra console ou choisir un rôle IAM existant à utiliser. La console affiche les rôles dont le nom contient la chaîne « kendra » ou « Kendra ».

Les rubriques suivantes fournissent des informations détaillées sur les politiques requises. Si vous créez des IAM rôles à l'aide de la Amazon Kendra console, ces politiques sont créées pour vous.

Rubriques

- IAM rôles pour les index
- IAM rôles pour l' BatchPutDocumentAPI
- IAM rôles pour les sources de données
- Rôle de cloud privé virtuel (VPC) IAM
- IAM rôles pour les questions fréquemment posées (FAQs)
- IAM rôles pour les suggestions de requêtes
- IAM rôles pour le mappage principal des utilisateurs et des groupes
- IAM rôles pour AWS IAM Identity Center
- IAM rôles liés aux Amazon Kendra expériences
- IAM rôles pour l'enrichissement de documents personnalisés

IAM rôles pour les index

Lorsque vous créez un index, vous devez fournir un IAM rôle autorisé à écrire dans un Amazon CloudWatch. Vous devez également fournir une politique de confiance qui permet Amazon Kendra d'assumer le rôle. Les politiques qui doivent être fournies sont les suivantes.

IAM rôles pour les index

Une politique de rôle permettant Amazon Kendra d'accéder à un CloudWatch journal.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": "cloudwatch:PutMetricData",
            "Resource": "*",
            "Condition": {
                "StringEquals": {
                    "cloudwatch:namespace": "AWS/Kendra"
                }
            }
        },
        {
            "Effect": "Allow",
            "Action": "logs:DescribeLogGroups",
            "Resource": "*"
        },
        {
            "Effect": "Allow",
            "Action": "logs:CreateLogGroup",
            "Resource": "arn:aws:logs:your-region:your-account-id:log-group:/aws/
kendra/*"
        },
        {
            "Effect": "Allow",
            "Action": [
                "logs:DescribeLogStreams",
                "logs:CreateLogStream",
                "logs:PutLogEvents"
            ],
            "Resource": "arn:aws:logs:your-region:your-account-id:log-group:/aws/
kendra/*:log-stream:*"
        }
    ]
}
```

Une politique de rôle pour Amazon Kendra autoriser l'accès AWS Secrets Manager. Si vous utilisez le contexte utilisateur Secrets Manager comme emplacement clé, vous pouvez appliquer la politique suivante.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Action":"cloudwatch:PutMetricData",
         "Resource":"*",
         "Condition":{
            "StringEquals":{
               "cloudwatch:namespace":"AWS/Kendra"
            }
         }
      },
      {
         "Effect":"Allow",
         "Action":"logs:DescribeLogGroups",
         "Resource":"*"
      },
      {
         "Effect":"Allow",
         "Action":"logs:CreateLogGroup",
         "Resource":"arn:aws:logs:your-region:your-account-id:log-group:/aws/
kendra/*"
      },
      {
         "Effect":"Allow",
         "Action":[
            "logs:DescribeLogStreams",
            "logs:CreateLogStream",
            "logs:PutLogEvents"
         ],
         "Resource":"arn:aws:logs:your-region:your-account-id:log-group:/aws/
kendra/*:log-stream:*"
      },
      {
         "Effect":"Allow",
```

```
            "Action":[
               "secretsmanager:GetSecretValue"
            ],
            "Resource":[
               "arn:aws:secretsmanager:your-region:your-account-id:secret:secret-id"
            ]
         },
         {
            "Effect":"Allow",
            "Action":[
               "kms:Decrypt"
            ],
            "Resource":[
               "arn:aws:kms:your-region:your-account-id:key/key-id"
            ],
            "Condition":{
               "StringLike":{
                  "kms:ViaService":[
                     "secretsmanager.your-region.amazonaws.com"
                  ]
               }
            }
         }
      ]
}
```

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
```

```
}
```

IAM rôles pour l' BatchPutDocumentAPI

> **Warning**
>
> Amazon Kendra n'utilise pas de politique de compartiment qui autorise un Amazon Kendra mandant à interagir avec un compartiment S3. Il utilise plutôt des IAM rôles. Assurez-vous qu'il Amazon Kendra n'est pas inclus en tant que membre de confiance dans votre politique de compartiment afin d'éviter tout problème de sécurité des données lié à l'octroi accidentel d'autorisations à des principaux arbitraires. Vous pouvez toutefois ajouter une politique de compartiment pour utiliser un Amazon S3 compartiment sur différents comptes. Pour plus d'informations, consultez la section Politiques applicables à Amazon S3 tous les comptes. Pour plus d'informations sur IAM les rôles pour les sources de données S3, consultez la section IAM rôles.

Lorsque vous utilisez l'BatchPutDocumentAPI pour indexer des documents dans un Amazon S3 compartiment, vous devez Amazon Kendra fournir un IAM rôle permettant d'accéder au compartiment. Vous devez également fournir une politique de confiance qui permet Amazon Kendra d'assumer le rôle. Si les documents du compartiment sont chiffrés, vous devez autoriser l'utilisation de la clé principale du AWS KMS client (CMK) pour les déchiffrer.

IAM rôles pour l' BatchPutDocumentAPI

Une politique de rôle obligatoire pour autoriser Amazon Kendra l'accès à un Amazon S3 compartiment.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "s3:GetObject"
```

```
            ],
            "Resource": [
                "arn:aws:s3:::bucket-name/*"
            ]
        }
    ]
}
```

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

Il est recommandé d'inclure `aws:sourceAccount` et `aws:sourceArn` dans la politique de confiance. Cela limite les autorisations et vérifie en toute sécurité si `aws:sourceAccount` et si `aws:sourceArn` elles sont identiques à celles prévues dans la politique de IAM rôle pour l'`sts:AssumeRole`action. Cela empêche les entités non autorisées d'accéder à vos IAM rôles et à leurs autorisations. Pour plus d'informations, consultez le AWS Identity and Access Management guide sur le problème de la confusion chez les députés.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
```

```
            "Effect": "Allow",
            "Principal": {
                "Service": [
                    "kendra.amazonaws.com"
                ]
            },
            "Action": "sts:AssumeRole",
            "Condition": {
                "StringEquals": {
                    "aws:SourceAccount": "your-account-id"
                },
                "StringLike": {
                    "aws:SourceArn": "arn:aws:kendra:your-region:your-account-
id:index/*"
                }
            }
        }
    ]
}
```

Politique de rôle facultative autorisant l'utilisation Amazon Kendra d'une clé principale AWS KMS du client (CMK) pour déchiffrer les documents d'un Amazon S3 compartiment.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "kms:Decrypt"
            ],
            "Resource": [
                "arn:aws:kms:your-region:your-account-id:key/key-id"
            ]
        }
    ]
}
```

IAM rôles pour les sources de données

Lorsque vous utilisez l'CreateDataSourceAPI, vous devez attribuer Amazon Kendra un IAM rôle autorisé à accéder aux ressources. Les autorisations spécifiques requises dépendent de la source de données.

IAM rôles pour les sources de données Adobe Experience Manager

Lorsque vous utilisez Adobe Experience Manager, vous attribuez un rôle soumis aux politiques suivantes.

- Autorisation d'accéder à votre AWS Secrets Manager code secret pour authentifier votre Adobe Experience Manager.
- Autorisation d'appeler le public requis APIs pour le connecteur Adobe Experience Manager.
- Autorisation d'appeler le `BatchPutDocument`,`BatchDeleteDocument`,`PutPrincipalMapping`,`DeletePrincipalMapping`,`D` et `ListGroupsOlderThanOrderingId` APIs.

> **Note**
>
> Vous pouvez connecter une source de données Adobe Experience Manager à Amazon Kendra via Amazon VPC. Si vous utilisez un Amazon VPC, vous devez ajouter des autorisations supplémentaires.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:{{your-region}}:{{your-account-id}}:secret:
[[secret-id]]"
```

```
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:{{your-region}}:{{your-account-id}}:key/[[key-id]]"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.{{your-region}}.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
    ],
    "Resource": ["arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/
{{index-id}}", "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{index-
id}}/data-source/*"]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/
{{index-id}}"
  }]
}
```

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM rôles pour les sources de données Alfresco

Lorsque vous utilisez Alfresco, vous attribuez un rôle avec les politiques suivantes.

- Autorisation d'accéder à votre AWS Secrets Manager secret pour authentifier votre Alfresco.
- Autorisation d'appeler le public requis APIs pour le connecteur Alfresco.
- Autorisation d'appeler le `BatchPutDocument`,`BatchDeleteDocument`,`PutPrincipalMapping`,`DeletePrincipalMapping`,`D` et `ListGroupsOlderThanOrderingId` APIs.

> **Note**
>
> Vous pouvez connecter une source de données Alfresco à Amazon Kendra via. Amazon VPC Si vous utilisez un Amazon VPC, vous devez ajouter des autorisations supplémentaires.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
```

```
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:{{your-region}}:{{your-account-id}}:secret:
[[secret-id]]"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:{{your-region}}:{{your-account-id}}:key/[[key-id]]"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.{{your-region}}.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
    ],
    "Resource": ["arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/
{{index-id}}", "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{index-
id}}/data-source/*"]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
```

```
    "Resource": "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/
{{index-id}}"
  }]
}
```

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM rôles pour les sources de données Aurora (MySQL)

Lorsque vous utilisez Aurora (MySQL), vous fournissez un rôle avec les politiques suivantes.

- Autorisation d'accéder à votre AWS Secrets Manager secret pour vous authentifier Aurora (MySQL).
- Autorisation d'appeler le public requis APIs pour le connecteur Aurora (MySQL).
- Autorisation d'appeler le `BatchPutDocument`,`BatchDeleteDocument`,`PutPrincipalMapping`,`DeletePrincipalMapping`,`D` et `ListGroupsOlderThanOrderingId` APIs.

> **Note**
>
> Vous pouvez connecter une source de données Aurora (MySQL) à Amazon Kendra through Amazon VPC. Si vous utilisez un Amazon VPC, vous devez ajouter des autorisations supplémentaires.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:{{region}}:{{account_id}}:secret:[[secret_id]]"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:{{region}}:{{account_id}}:key/[[key_id]]"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.*.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
```

```
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
    ],
    "Resource": ["arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}",
 "arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}/data-source/*"]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}"
  }]
}
```

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM rôles pour les Aurora sources de données (PostgreSQL)

Lorsque vous utilisez Aurora (PostgreSQL), vous fournissez un rôle avec les politiques suivantes.

- Autorisation d'accéder à votre AWS Secrets Manager code secret pour vous authentifier Aurora (PostgreSQL).

- Autorisation d'appeler le public requis APIs pour le connecteur Aurora (PostgreSQL).
- Autorisation d'appeler le `BatchPutDocument`,`BatchDeleteDocument`,`PutPrincipalMapping`,`DeletePrincipalMapping`,`D` et `ListGroupsOlderThanOrderingId` APIs.

> **Note**
>
> Vous pouvez connecter une source de données Aurora (PostgreSQL) à via. Amazon Kendra Amazon VPC Si vous utilisez un Amazon VPC, vous devez ajouter des autorisations supplémentaires.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:{{region}}:{{account_id}}:secret:[[secret_id]]"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:{{region}}:{{account_id}}:key/[[key_id]]"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.*.amazonaws.com"
        ]
      }
```

```
    }
  },
  {
    "Effect": "Allow",
    "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
    ],
    "Resource": ["arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}",
 "arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}/data-source/*"]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}"
  }]
}
```

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM rôles pour les sources Amazon FSx de données

Lorsque vous utilisez Amazon FSx, vous fournissez un rôle avec les politiques suivantes.

- Autorisation d'accéder à votre AWS Secrets Manager secret pour authentifier votre système de Amazon FSx fichiers.
- Autorisation d'accès Amazon Virtual Private Cloud (VPC) à l'endroit où réside votre système de Amazon FSx fichiers.
- Autorisation d'obtenir le nom de domaine de votre Active Directory pour votre système de Amazon FSx fichiers.
- Autorisation d'appeler le public requis APIs pour le Amazon FSx connecteur.
- Autorisation d'appeler le `BatchPutDocument` et `BatchDeleteDocument` APIs de mettre à jour l'index.

JSON

```
{
    "Version": "2012-10-17",
        "Statement": [
        {
          "Effect": "Allow",
          "Action": [
            "secretsmanager:GetSecretValue"
          ],
          "Resource": [
            "arn:aws:secretsmanager:{{your-region}}:{{your-account-id}}:secret:
{{secret-id}}"
          ]
        },
        {
          "Effect": "Allow",
          "Action": [
            "kms:Decrypt"
          ],
          "Resource": [
            "arn:aws:kms:{{your-region}}:{{your-account-id}}:key/{{key-id}}"
          ],
          "Condition": {
            "StringLike": {
```

```
            "kms:ViaService": [
              "secretsmanager.{{your-region}}.amazonaws.com"
            ]
          }
        }
      },
      {
        "Effect": "Allow",
        "Action":[
          "ec2:CreateNetworkInterface",
          "ec2:DeleteNetworkInterface"
        ],
        "Resource": [
             "arn:aws:ec2:{{your-region}}:{{your-account-id}}:network-
interface/*",
             "arn:aws:ec2:{{your-region}}:{{your-account-id}}:subnet/[[subnet-
ids]]"
        ]
      },
      {
        "Effect": "Allow",
        "Action": [
          "ec2:DescribeSubnets",
          "ec2:DescribeNetworkInterfaces"
        ],
        "Resource": "*"
      },
      {
        "Effect": "Allow",
        "Action": [
          "ec2:CreateNetworkInterfacePermission"
        ],
        "Resource": "arn:aws:ec2:{{your-region}}:{{your-account-id}}:network-
interface/*",
        "Condition": {
          "StringEquals": {
            "ec2:AuthorizedService": "kendra.*.amazonaws.com"
          },
          "ArnEquals": {
            "ec2:Subnet": [
              "arn:aws:ec2:{{your-region}}:{{your-account-id}}:subnet/[[subnet-
ids]]"
            ]
          }
```

```
          }
        },
        {
          "Sid": "AllowsKendraToGetDomainNameOfActiveDirectory",
          "Effect": "Allow",
          "Action": "ds:DescribeDirectories",
          "Resource": "*"
        },
        {
          "Sid": "AllowsKendraToCallRequiredFsxAPIs",
          "Effect": "Allow",
          "Action": [
              "fsx:DescribeFileSystems"
          ],
          "Resource": "*"
        },
        {
          "Sid": "iamPassRole",
          "Effect": "Allow",
          "Action": "iam:PassRole",
          "Resource": "*",
          "Condition": {
            "StringEquals": {
              "iam:PassedToService": [
                "kendra.*.amazonaws.com"
              ]
            }
          }
        },
        {
          "Effect": "Allow",
          "Action": [
            "kendra:BatchPutDocument",
            "kendra:BatchDeleteDocument"
          ],
          "Resource": "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/
{{index-id}}"
        }
        ]
}
```

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM rôles pour les sources de données de base de données

Lorsque vous utilisez une base de données comme source de données, vous Amazon Kendra attribuez un rôle doté des autorisations nécessaires pour se connecter au. Il s'agit des licences suivantes :

- Autorisation d'accéder au AWS Secrets Manager secret qui contient le nom d'utilisateur et le mot de passe du site. Pour plus d'informations sur le contenu du secret, consultez la section Sources de données.
- Autorisation d'utiliser la clé principale du AWS KMS client (CMK) pour déchiffrer le nom d'utilisateur et le secret du mot de passe stockés par. Secrets Manager
- Autorisation d'utiliser les `BatchDeleteDocument` opérations `BatchPutDocument` et pour mettre à jour l'index.
- Autorisation d'accéder au Amazon S3 compartiment contenant le certificat SSL utilisé pour communiquer avec le site.

> **Note**
>
> Vous pouvez connecter des sources de données de base de données Amazon Kendra via Amazon VPC. Si vous utilisez un Amazon VPC, vous devez ajouter des autorisations supplémentaires.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "secretsmanager:GetSecretValue"
            ],
            "Resource": [
                "arn:aws:secretsmanager:your-region:your-account-id:secret:secret-id"
            ]
        },
        {
            "Effect": "Allow",
            "Action": [
                "kms:Decrypt"
            ],
            "Resource": [
                "arn:aws:kms:your-region:your-account-id:key/key-id"
            ]
        },
        {
            "Effect": "Allow",
            "Action": [
                "kendra:BatchPutDocument",
                "kendra:BatchDeleteDocument"
            ],
            "Resource": [
                "arn:aws:kendra:your-region:your-account-id:index/index-id"
            "Condition": {
                "StringLike": {
                    "kms:ViaService": [
                        "kendra.your-region.amazonaws.com"
                    ]
                }
            }
        },
        {
            "Effect": "Allow",
            "Action": [
                "s3:GetObject"
            ],
            "Resource": [
```

```
                "arn:aws:s3:::bucket-name/*"
            ]
        }
    ]
}
```

Il existe deux politiques facultatives que vous pouvez utiliser avec une source de données.

Si vous avez chiffré le Amazon S3 compartiment contenant le certificat SSL utilisé pour communiquer avec le, fournissez une politique pour donner Amazon Kendra accès à la clé.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "kms:Decrypt"
            ],
            "Resource": [
                "arn:aws:kms:your-region:your-account-id:key/key-id"
            ]
        }
    ]
}
```

Si vous utilisez un VPC, définissez une politique qui donne Amazon Kendra accès aux ressources requises. Voir IAM les rôles pour les sources de données, VPC pour la politique requise.

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
```

```
            "Principal":{
               "Service":"kendra.amazonaws.com"
            },
            "Action":"sts:AssumeRole"
         }
      ]
}
```

IAM rôles pour les sources de données Amazon RDS (Microsoft SQL Server)

Lorsque vous utilisez un connecteur de source de données Amazon RDS (Microsoft SQL Server), vous fournissez un rôle avec les politiques suivantes.

- Autorisation d'accéder à votre code AWS Secrets Manager secret pour authentifier votre instance de source de données Amazon RDS (Microsoft SQL Server).
- Autorisation d'appeler le public requis APIs pour le connecteur de source de données Amazon RDS (Microsoft SQL Server).
- Autorisation d'appeler le `BatchPutDocument`,`BatchDeleteDocument`,`PutPrincipalMapping`,`DeletePrincipalMapping`,`D` et `ListGroupsOlderThanOrderingId` APIs.

> **Note**
>
> Vous pouvez connecter une source de données Amazon RDS (Microsoft SQL Server) à Amazon Kendra via Amazon VPC. Si vous utilisez un Amazon VPC, vous devez ajouter des autorisations supplémentaires.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
```

```
    "Resource": [
      "arn:aws:secretsmanager:{{region}}:{{account_id}}:secret:[[secret_id]]"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:{{region}}:{{account_id}}:key/[[key_id]]"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.*.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
    ],
    "Resource": ["arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}",
 "arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}/data-source/*"]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}"
  }]
}
```

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM rôles pour les sources de données Amazon RDS (MySQL)

Lorsque vous utilisez un connecteur de source de données Amazon RDS (MySQL), vous fournissez un rôle avec les politiques suivantes.

- Autorisation d'accéder à votre AWS Secrets Manager secret pour authentifier votre instance de source de données Amazon RDS (MySQL).
- Autorisation d'appeler le public requis APIs pour le connecteur de source de données Amazon RDS (MySQL).
- Autorisation d'appeler le `BatchPutDocument`,`BatchDeleteDocument`,`PutPrincipalMapping`,`DeletePrincipalMapping`,`D` et `ListGroupsOlderThanOrderingId` APIs.

> **Note**
>
> Vous pouvez connecter une source de données Amazon RDS (MySQL) à Amazon Kendra through Amazon VPC. Si vous utilisez un Amazon VPC, vous devez ajouter des autorisations supplémentaires.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:{{region}}:{{account_id}}:secret:[[secret_id]]"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:{{region}}:{{account_id}}:key/[[key_id]]"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.*.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
    ],
    "Resource": ["arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}",
 "arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}/data-source/*"]
  },
  {
    "Effect": "Allow",
```

```
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}"
  }]
}
```

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM rôles pour les sources de données Amazon RDS (Oracle)

Lorsque vous utilisez un connecteur de source de données Amazon RDS Oracle, vous fournissez un rôle avec les politiques suivantes.

- Autorisation d'accéder à votre code AWS Secrets Manager secret pour authentifier votre instance de source de données Amazon RDS (Oracle).
- Autorisation d'appeler le public requis APIs pour le connecteur de source de données Amazon RDS (Oracle).
- Autorisation d'appeler le `BatchPutDocumentBatchDeleteDocument`,`PutPrincipalMapping`,`DeletePrincipalMapping`,`D` et `ListGroupsOlderThanOrderingId` APIs.

> **Note**
>
> Vous pouvez connecter une source de données Amazon RDS Oracle à Amazon Kendra via Amazon VPC. Si vous utilisez un Amazon VPC, vous devez ajouter des autorisations supplémentaires.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:{{region}}:{{account_id}}:secret:[[secret_id]]"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:{{region}}:{{account_id}}:key/[[key_id]]"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.*.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
```

```
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
    ],
    "Resource": ["arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}",
 "arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}/data-source/*"]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}"
  }]
}
```

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM rôles pour les Amazon RDS sources de données (PostgreSQL)

Lorsque vous utilisez un connecteur de source de données Amazon RDS (PostgreSQL), vous fournissez un rôle avec les politiques suivantes.

- Autorisation d'accéder à votre AWS Secrets Manager secret pour authentifier votre instance de source de données Amazon RDS (PostgreSQL).
- Autorisation d'appeler le public requis APIs pour le connecteur de Amazon RDS source de données (PostgreSQL).
- Autorisation d'appeler le `BatchPutDocument`,`BatchDeleteDocument`,`PutPrincipalMapping`,`DeletePrincipalMapping`,`D` et `ListGroupsOlderThanOrderingId` APIs.

> **Note**
>
> Vous pouvez connecter une source de données Amazon RDS (PostgreSQL) à via. Amazon Kendra Amazon VPC Si vous utilisez un Amazon VPC, vous devez ajouter des autorisations supplémentaires.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:{{region}}:{{account_id}}:secret:[[secret_id]]"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:{{region}}:{{account_id}}:key/[[key_id]]"
    ],
    "Condition": {
      "StringLike": {
```

```
        "kms:ViaService": [
          "secretsmanager.*.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
    ],
    "Resource": ["arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}",
 "arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}/data-source/*"]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}"
  }]
}
```

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
```

```
    ]
}
```

IAM rôles pour les sources Amazon S3 de données

> **Warning**
>
> Amazon Kendra n'utilise pas de politique de compartiment qui autorise un Amazon Kendra mandant à interagir avec un compartiment S3. Il utilise plutôt des IAM rôles. Assurez-vous qu'il Amazon Kendra n'est pas inclus en tant que membre de confiance dans votre politique de compartiment afin d'éviter tout problème de sécurité des données lié à l'octroi accidentel d'autorisations à des principaux arbitraires. Vous pouvez toutefois ajouter une politique de compartiment pour utiliser un Amazon S3 compartiment sur différents comptes. Pour plus d'informations, voir Politiques à utiliser Amazon S3 sur tous les comptes (faites défiler l'écran vers le bas).

Lorsque vous utilisez un Amazon S3 bucket comme source de données, vous fournissez un rôle autorisé à accéder au bucket et à utiliser les `BatchDeleteDocument` opérations `BatchPutDocument` et. Si les documents du Amazon S3 compartiment sont chiffrés, vous devez autoriser l'utilisation de la clé principale du AWS KMS client (CMK) pour les déchiffrer.

Les politiques de rôle suivantes doivent Amazon Kendra permettre d'assumer un rôle. Faites défiler la page vers le bas pour voir une politique de confiance permettant d'assumer un rôle.

Une politique de rôle obligatoire pour Amazon Kendra autoriser l'utilisation d'un Amazon S3 bucket comme source de données.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Action": [
                "s3:GetObject"
            ],
            "Resource": [
```

```
                "arn:aws:s3:::bucket-name/*"
            ],
            "Effect": "Allow"
        },
        {
            "Action": [
                "s3:ListBucket"
            ],
            "Resource": [
                "arn:aws:s3:::bucket-name"
            ],
            "Effect": "Allow"
        },
        {
            "Effect": "Allow",
            "Action": [
                "kendra:BatchPutDocument",
                "kendra:BatchDeleteDocument"
            ],
            "Resource": [
                "arn:aws:kendra:your-region:your-account-id:index/index-id"
            ]
        }
    ]
}
```

Politique de rôle facultative autorisant l'utilisation Amazon Kendra d'une clé principale AWS KMS du client (CMK) pour déchiffrer les documents d'un Amazon S3 compartiment.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "kms:Decrypt"
            ],
            "Resource": [
                "arn:aws:kms:your-region:your-account-id:key/key-id"
```

```
            ]
        }
    ]
}
```

Une politique de rôle facultative permettant d'accéder Amazon Kendra à un Amazon S3 compartiment, tout en utilisant un Amazon VPC, et sans activer AWS KMS ni partager AWS KMS d'autorisations.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Action": [
        "s3:GetObject"
      ],
      "Resource": [
        "arn:aws:s3:::{{bucket-name}}/*"
      ],
      "Effect": "Allow"
    },
    {
      "Action": [
        "s3:ListBucket"
      ],
      "Resource": [
        "arn:aws:s3:::{{bucket-name}}"
      ],
      "Effect": "Allow"
    },
    {
      "Effect": "Allow",
      "Action": [
        "ec2:CreateNetworkInterface"
      ],
      "Resource": [
        "arn:aws:ec2:{{your-region}}:{{your-account-id}}:subnet/[[subnet-ids]]",
        "arn:aws:ec2:{{your-region}}:{{your-account-id}}:security-group/
[[security-group]]"
```

```
      ]
    },
    {
        "Effect": "Allow",
        "Action": [
            "ec2:CreateNetworkInterface"
        ],
        "Resource": "arn:aws:ec2:{{your-region}}:{{your-account-id}}:network-
interface/*",
        "Condition": {
            "StringLike": {
                "aws:RequestTag/AWS_KENDRA": "kendra_{{your-account-id}}_{{index-
id}}_{data-source-id}}_*"
            }
        }
    },
    {
        "Effect": "Allow",
        "Action": [
            "ec2:CreateTags"
        ],
        "Resource": "arn:aws:ec2:{{your-region}}:{{your-account-id}}:network-
interface/*",
        "Condition": {
            "StringEquals": {
                "ec2:CreateAction": "CreateNetworkInterface"
            }
        }
    },
    {
      "Effect": "Allow",
      "Action": [
        "ec2:DescribeSubnets"
      ],
      "Resource": "*"
    },
    {
      "Effect": "Allow",
      "Action": [
        "ec2:DescribeNetworkInterfaces"
      ],
      "Resource": "*"
    },
    {
```

```
      "Effect": "Allow",
      "Action": [
        "ec2:CreateNetworkInterfacePermission"
      ],
      "Resource": "arn:aws:ec2:{{your-region}}:{{your-accoount-id}}:network-
interface/*",
      "Condition": {
        "StringEquals": {
          "ec2:AuthorizedService": "kendra.amazonaws.com"
        },
        "ArnEquals": {
          "ec2:Subnet": [
            "arn:aws:ec2:{{your-region}}:{{your-account-id}}:subnet/[[subnet-
ids]]"
          ]
        }
      }
    },
    {
      "Effect": "Allow",
      "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
      ],
      "Resource": [
        "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{index-id}}",
        "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{index-id}}/
data-source/*"
      ]
    },
    {
      "Effect": "Allow",
      "Action": [
        "kendra:BatchPutDocument",
        "kendra:BatchDeleteDocument"
      ],
      "Resource": "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/
{{index-id}}"
    }
  ]
}
```

Une politique de rôle facultative permettant d'accéder Amazon Kendra à un Amazon S3 compartiment en utilisant un Amazon VPC et avec AWS KMS les autorisations activées.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Action": [
        "s3:GetObject"
      ],
      "Resource": [
        "arn:aws:s3:::{{bucket-name}}/*"
      ],
      "Effect": "Allow"
    },
    {
      "Action": [
        "s3:ListBucket"
      ],
      "Resource": [
        "arn:aws:s3:::{{bucket-name}}"
      ],
      "Effect": "Allow"
    },
    {
      "Effect": "Allow",
      "Action": [
        "kms:Decrypt"
      ],
      "Resource": [
        "arn:aws:kms:{{your-region}}:{{your-account-id}}:key/{{key-id}}"
      ],
      "Condition": {
        "StringLike": {
          "kms:ViaService": [
            "s3.{{your-region}}.amazonaws.com"
          ]
        }
      }
    },
    {
```

```
      "Effect": "Allow",
      "Action": [
        "ec2:CreateNetworkInterface"
      ],
      "Resource": [
        "arn:aws:ec2:{{your-region}}:{{your-account-id}}:subnet/[[subnet-ids]]",
        "arn:aws:ec2:{{your-region}}:{{your-account-id}}:security-group/
[[security-group]]"
      ]
    },
    {
        "Effect": "Allow",
        "Action": [
            "ec2:CreateNetworkInterface"
        ],
        "Resource": "arn:aws:ec2:{{your-region}}:{{your-account-id}}:network-
interface/*",
        "Condition": {
            "StringLike": {
                "aws:RequestTag/AWS_KENDRA": "kendra_{{your-account-id}}_{{index-
id}}_{data-source-id}}_*"
            }
        }
    },
    {
        "Effect": "Allow",
        "Action": [
            "ec2:CreateTags"
        ],
        "Resource": "arn:aws:ec2:{{your-region}}:{{your-account-id}}:network-
interface/*",
        "Condition": {
            "StringEquals": {
                "ec2:CreateAction": "CreateNetworkInterface"
            }
        }
    },
    {
      "Effect": "Allow",
      "Action": [
        "ec2:DescribeSubnets"
      ],
      "Resource": "*"
    },
```

```
    {
      "Effect": "Allow",
      "Action": [
        "ec2:DescribeNetworkInterfaces"
      ],
      "Resource": "*"
    },
    {
      "Effect": "Allow",
      "Action": [
        "ec2:CreateNetworkInterfacePermission"
      ],
      "Resource": "arn:aws:ec2:{{your-region}}:{{your-account-id}}:network-
interface/*",
      "Condition": {
        "StringEquals": {
          "ec2:AuthorizedService": "kendra.amazonaws.com"
        },
        "ArnEquals": {
          "ec2:Subnet": [
            "arn:aws:ec2:{{your-region}}:{{your-account-id}}:subnet/[[subnet-
ids]]"
          ]
        }
      }
    },
    {
      "Effect": "Allow",
      "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
      ],
      "Resource": [
        "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{index-id}}",
        "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{index-id}}/
data-source/*"
      ]
    },
    {
      "Effect": "Allow",
      "Action": [
        "kendra:BatchPutDocument",
```

```
        "kendra:BatchDeleteDocument"
      ],
      "Resource": "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/
{{index-id}}"
    }
  ]
}
```

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

Politiques à utiliser Amazon S3 sur tous les comptes

Si votre Amazon S3 bucket se trouve dans un compte différent de celui que vous utilisez pour votre Amazon Kendra index, vous pouvez créer des règles pour l'utiliser sur tous les comptes.

Une politique de rôle permettant d'utiliser votre Amazon S3 bucket comme source de données lorsque le bucket se trouve dans un compte différent de celui de votre Amazon Kendra index. Notez que `s3:PutObject` et `s3:PutObjectAcl` sont facultatifs, et vous pouvez les utiliser si vous souhaitez inclure un fichier de configuration pour votre liste de contrôle d'accès.

JSON

```
{
```

```
    "Version": "2012-10-17",
    "Statement": [
         {
            "Action": [
                "s3:GetObject"
            ],
            "Resource": [
                "arn:aws:s3:::$bucket-in-other-account/*"
            ],
            "Effect": "Allow"
        },
        {
            "Action": [
                "s3:ListBucket"
            ],
            "Resource": [
                "arn:aws:s3:::$bucket-in-other-account/*"
            ],
            "Effect": "Allow"
        },
        {
            "Effect": "Allow",
            "Action": [
                "kendra:BatchPutDocument",
                "kendra:BatchDeleteDocument"
            ],
            "Resource": [
                "arn:aws:kendra:$your-region:$your-account-id:index/$index-id"
            ]
        },
        {
            "Effect": "Allow",
            "Action": [
                "s3:GetObject",
                "s3:PutObject",
                "s3:PutObjectAcl"
            ],
            "Resource": "arn:aws:s3:::$bucket-in-other-account/*"
        }
    ]
}
```

Une politique de compartiment permettant au rôle de source de Amazon S3 données d'accéder au Amazon S3 compartiment entre les comptes. Notez que `s3:PutObject` et `s3:PutObjectAcl` sont facultatifs, et vous pouvez les utiliser si vous souhaitez inclure un fichier de configuration pour votre liste de contrôle d'accès.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Principal": {
                "AWS": "$kendra-s3-connector-role-arn"
            },
            "Action": [
                "s3:GetObject",
                "s3:PutObject",
                "s3:PutObjectAcl"
            ],
            "Resource": [
                "arn:aws:s3:::$bucket-in-other-account/*"
            ]
        },
        {
            "Effect": "Allow",
            "Principal": {
                "AWS": "$kendra-s3-connector-role-arn"
            },
            "Action": "s3:ListBucket",
            "Resource": "arn:aws:s3:::$bucket-in-other-account"
        }
    ]
}
```

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM rôles pour les sources de données Amazon Kendra Web Crawler

Lorsque vous utilisez Amazon Kendra Web Crawler, vous attribuez un rôle avec les politiques suivantes :

- Autorisation d'accéder au AWS Secrets Manager secret qui contient les informations d'identification pour se connecter à des sites Web ou à un serveur proxy Web soutenu par une authentification de base. Pour plus d'informations sur le contenu du secret, consultez la section Utilisation d'une source de données de robot d'exploration Web.
- Autorisation d'utiliser la clé principale du AWS KMS client (CMK) pour déchiffrer le nom d'utilisateur et le secret du mot de passe stockés par. Secrets Manager
- Autorisation d'utiliser les `BatchDeleteDocument` opérations `BatchPutDocument` et pour mettre à jour l'index.
- Si vous utilisez un Amazon S3 bucket pour stocker votre liste de graines URLs ou vos plans de site, incluez l'autorisation d'accéder au Amazon S3 bucket.

> **Note**
>
> Vous pouvez connecter une source de données Amazon Kendra Web Crawler à Amazon Kendra via Amazon VPC. Si vous utilisez un Amazon VPC, vous devez ajouter des autorisations supplémentaires.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:your-region:your-account-id:secret:secret-id"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:your-region:your-account:key/key-id"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.your-region.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:your-region:your-account-id:index/index-id"
  }]
}
```

Si vous stockez vos graines URLs ou vos plans de site dans un Amazon S3 bucket, vous devez ajouter cette autorisation au rôle.

```
,
{"Effect": "Allow",
     "Action": [
         "s3:GetObject"
      ],
      "Resource": [
         "arn:aws:s3:::bucket-name/*"
      ]
}
```

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM rôles pour les sources de données Box

Lorsque vous utilisez Box, vous attribuez un rôle avec les politiques suivantes.

- Autorisation d'accéder à votre AWS Secrets Manager secret pour authentifier votre Slack.
- Autorisation d'appeler le public requis APIs pour le connecteur Box.
- Autorisation d'appeler le `BatchPutDocument`,`BatchDeleteDocument`,`PutPrincipalMapping`,`DeletePrincipalMapping`,`D` et `ListGroupsOlderThanOrderingId` APIs.

> **Note**
>
> Vous pouvez connecter une source de données Box à Amazon Kendra via Amazon VPC. Si vous utilisez un Amazon VPC, vous devez ajouter des autorisations supplémentaires.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:{{your-region}}:{{your-account-id}}:secret:
[[secret-id]]"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:{{your-region}}:{{your-account-id}}:key/[[key-id]]"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.{{your-region}}.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
```

```
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
    ],
    "Resource": ["arn:aws:kendra:{{your-region}}:{{your-account-d}}:index/
{{index-id}}", "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{index-
id}}/data-source/*"]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/
{{index-id}}"
  }]
}
```

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM rôles pour les sources de données Confluence

IAM rôles pour Confluence Connector v1.0

Lorsque vous utilisez un serveur Confluence comme source de données, vous fournissez un rôle avec les politiques suivantes :

- Autorisation d'accéder au AWS Secrets Manager secret contenant les informations d'identification nécessaires pour se connecter à Confluence. Pour plus d'informations sur le contenu du secret, consultez la section Sources de données Confluence.
- Autorisation d'utiliser la clé principale du AWS KMS client (CMK) pour déchiffrer le nom d'utilisateur et le secret du mot de passe stockés par. Secrets Manager
- Autorisation d'utiliser les `BatchDeleteDocument` opérations `BatchPutDocument` et pour mettre à jour l'index.

> **Note**
>
> Vous pouvez connecter une source de données Confluence à Amazon Kendra via Amazon VPC. Si vous utilisez un Amazon VPC, vous devez ajouter des autorisations supplémentaires.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:your-region:your-account-id:secret:secret-id"
    ]
  },
  {
    "Effect": "Allow",
```

```
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:your-region:your-account-id:key/key-id"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.your-region.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:your-region:your-account-id:index/index-id"
  }]
}
```

Si vous utilisez un VPC, définissez une politique qui donne Amazon Kendra accès aux ressources requises. Voir IAM les rôles pour les sources de données, VPC pour la politique requise.

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
```

```
}
```

IAM rôles pour Confluence Connector v2.0

Pour une source de données Confluence Connector v2.0, vous fournissez un rôle avec les politiques suivantes.

- Autorisation d'accéder au AWS Secrets Manager secret qui contient les informations d'authentification pour Confluence. Pour plus d'informations sur le contenu du secret, consultez la section Sources de données Confluence.
- Autorisation d'utiliser la clé principale du AWS KMS client (CMK) pour déchiffrer le nom d'utilisateur et le secret du mot de passe stockés par. AWS Secrets Manager
- Autorisation d'utiliser les `BatchDeleteDocument` opérations `BatchPutDocument` et pour mettre à jour l'index.

Vous devez également joindre une politique de confiance qui permet Amazon Kendra d'assumer le rôle.

> **Note**
>
> Vous pouvez connecter une source de données Confluence à Amazon Kendra via Amazon VPC. Si vous utilisez un Amazon VPC, vous devez ajouter des autorisations supplémentaires.

Une politique de rôle permettant de Amazon Kendra se connecter à Confluence.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "secretsmanager:GetSecretValue"
      ],
      "Resource": [
        "arn:aws:secretsmanager:your-region:your-account-id:secret:secret-id"
      ]
```

```
    },
    {
      "Effect": "Allow",
      "Action": [
        "kms:Decrypt"
      ],
      "Resource": [
        "arn:aws:kms:your-region:your-account-id:key/key-id"
      ],
      "Condition": {
        "StringLike": {
          "kms:ViaService": [
            "secretsmanager.your-region.amazonaws.com"
          ]
        }
      }
    },
    {
      "Effect": "Allow",
      "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
      ],
      "Resource": [
        "arn:aws:kendra:your-region:your-account-id:index/index-id",
        "arn:aws:kendra:your-region:your-account-id:index/index-id/data-source/*"
      ]
    }
    {
      "Effect": "Allow",
      "Action": [
        "kendra:BatchPutDocument",
        "kendra:BatchDeleteDocument"
      ],
      "Resource": "arn:aws:kendra:your-region:your-account-id:index/index-id"
    }
  ]
}
```

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM rôles pour les sources de données Dropbox

Lorsque vous utilisez Dropbox, vous attribuez un rôle conformément aux règles suivantes.

- Autorisation d'accéder à votre AWS Secrets Manager code secret pour authentifier votre Dropbox.
- Autorisation d'appeler le public requis APIs pour le connecteur Dropbox.
- Autorisation d'appeler le `BatchPutDocument`,`BatchDeleteDocument`,`PutPrincipalMapping`,`DeletePrincipalMapping`,`D` et `ListGroupsOlderThanOrderingId` APIs.

> **Note**
>
> Vous pouvez y connecter une source de données Dropbox Amazon Kendra via Amazon VPC. Si vous utilisez un Amazon VPC, vous devez ajouter des autorisations supplémentaires.

JSON

```
{
"Version": "2012-10-17",
  "Statement": [
```

```
  {"Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:{{your-region}}:{{your-account-id}}:secret:
[[secret-id]]"
    ]
  },
  {"Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:{{your-region}}:{{your-account-id}}:key/[[key-id]]"
    ],
    "Condition": {"StringLike": {"kms:ViaService": [
          "secretsmanager.{{your-region}}.amazonaws.com"
        ]
      }
    }
  },
  {"Effect": "Allow",
    "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
    ],
    "Resource": ["arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/
{{index-id}}", "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{index-
id}}/data-source/*"]
  },
  {"Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/
{{index-id}}"
  }]
}
```

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM rôles pour les sources de données Drupal

Lorsque vous utilisez Drupal, vous fournissez un rôle avec les politiques suivantes.

- Autorisation d'accéder à votre AWS Secrets Manager secret pour authentifier votre Drupal.
- Autorisation d'appeler le public requis APIs pour le connecteur Drupal.
- Autorisation d'appeler le `BatchPutDocument`,`BatchDeleteDocument`,`PutPrincipalMapping`,`DeletePrincipalMapping`,`D` et `ListGroupsOlderThanOrderingId` APIs.

> **Note**
>
> Vous pouvez connecter une source de données Drupal à Amazon Kendra via. Amazon VPC Si vous utilisez un Amazon VPC, vous devez ajouter des autorisations supplémentaires.

JSON

```
{
  "Version": "2012-10-17",
```

```
 "Statement": [
 {
   "Effect": "Allow",
   "Action": [
     "secretsmanager:GetSecretValue"
   ],
   "Resource": [
     "arn:aws:secretsmanager:{{region}}:{{account_id}}:secret:[[secret_id]]"
   ]
 },
 {
   "Effect": "Allow",
   "Action": [
     "kms:Decrypt"
   ],
   "Resource": [
     "arn:aws:kms:{{region}}:{{account_id}}:key/[[key_id]]"
   ],
   "Condition": {
     "StringLike": {
       "kms:ViaService": [
         "secretsmanager.*.amazonaws.com"
       ]
     }
   }
 },
 {
   "Effect": "Allow",
   "Action": [
       "kendra:PutPrincipalMapping",
       "kendra:DeletePrincipalMapping",
       "kendra:ListGroupsOlderThanOrderingId",
       "kendra:DescribePrincipalMapping"
   ],
   "Resource": ["arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}",
"arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}/data-source/*"]
 },
 {
   "Effect": "Allow",
   "Action": [
     "kendra:BatchPutDocument",
     "kendra:BatchDeleteDocument"
   ],
   "Resource": "arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}"
```

```
  }]
}
```

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM rôles pour les sources GitHub de données

Lorsque vous utilisez GitHub, vous fournissez un rôle avec les politiques suivantes.

- Autorisation d'accéder à votre AWS Secrets Manager secret pour authentifier votre GitHub.
- Autorisation d'appeler le public requis APIs pour le GitHub connecteur.
- Autorisation d'appeler le `BatchPutDocument`,`BatchDeleteDocument`,`PutPrincipalMapping`,`DeletePrincipalMapping`,`D` et `ListGroupsOlderThanOrderingId` APIs.

> **Note**
>
> Vous pouvez connecter une source de GitHub données Amazon Kendra via Amazon VPC. Si vous utilisez un Amazon VPC, vous devez ajouter des autorisations supplémentaires.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:{{your-region}}:{{your-account-id}}:secret:
[[secret-id]]"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:{{your-region}}:{{your-account-id}}:key/[[key-id]]"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.{{your-region}}.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
    ],
    "Resource": ["arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/
{{index-id}}", "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{index-
id}}/data-source/*"]
  },
```

```
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/
{{index-id}}"
  }]
}
```

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM rôles pour les sources de données Gmail

Lorsque vous utilisez Gmail, vous attribuez un rôle soumis aux règles suivantes.

- Autorisation d'accéder à votre AWS Secrets Manager secret pour authentifier votre compte Gmail.
- Autorisation d'appeler le public requis APIs pour le Gmailconnector.
- Autorisation d'appeler le `BatchPutDocument`,`BatchDeleteDocument`,`PutPrincipalMapping`,`DeletePrincipalMapping`,`D` et `ListGroupsOlderThanOrderingId` APIs.

> **Note**
>
> Vous pouvez connecter une source de données Gmail à Amazon Kendra via Amazon VPC. Si vous utilisez un Amazon VPC, vous devez ajouter des autorisations supplémentaires.

JSON

```
{
"Version": "2012-10-17",
  "Statement": [
  {"Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:{{your-region}}:{{your-account-id}}:secret:
[[secret-id]]"
    ]
  },
  {"Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:{{your-region}}:{{your-account-id}}:key/[[key-id]]"
    ],
    "Condition": {"StringLike": {"kms:ViaService": [
          "secretsmanager.{{your-region}}.amazonaws.com"
        ]
      }
    }
  },
  {"Effect": "Allow",
    "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
    ],
```

```
    "Resource": ["arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/
{{index-id}}", "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{index-
id}}/data-source/*"]
  },
  {"Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/
{{index-id}}"
  }]
}
```

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM rôles pour les sources de données Google Drive

Lorsque vous utilisez une source de données Google Workspace Drive, vous Amazon Kendra attribuez un rôle doté des autorisations nécessaires pour se connecter au site. Il s’agit des licences suivantes :

- Autorisation d'obtenir et de déchiffrer le AWS Secrets Manager secret qui contient l'adresse e-mail du compte client, l'adresse e-mail du compte administrateur et la clé privée nécessaires pour se

connecter au site Google Drive. Pour plus d'informations sur le contenu du secret, consultez la section Sources de données Google Drive.

- Autorisation d'utiliser le BatchPutDocumentet BatchDeleteDocument APIs.

> **Note**
>
> Vous pouvez connecter une source de données Google Drive à Amazon Kendra via Amazon VPC. Si vous utilisez un Amazon VPC, vous devez ajouter des autorisations supplémentaires.

La IAM politique suivante fournit les autorisations nécessaires :

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:your-region:your-account-id:secret:secret-id"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:your-region:your-account-id:key/key-id"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.your-region.amazonaws.com"
        ]
```

```
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:your-region:your-account-id:index/index-id"
  }]
}
```

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM rôles pour les sources DB2 de données IBM

Lorsque vous utilisez un connecteur de source de DB2 données IBM, vous fournissez un rôle avec les politiques suivantes.

- Autorisation d'accéder à votre AWS Secrets Manager secret pour authentifier votre instance de source de DB2 données IBM.
- Autorisation d'appeler le public requis APIs pour le connecteur de source de DB2 données IBM.

- Autorisation d'appeler le `BatchPutDocument`,`BatchDeleteDocument`,`PutPrincipalMapping`,`DeletePrincipalMapping`,`D` et `ListGroupsOlderThanOrderingId` APIs.

> **Note**
>
> Vous pouvez connecter une source de DB2 données IBM à Amazon Kendra via Amazon VPC. Si vous utilisez un Amazon VPC, vous devez ajouter des autorisations supplémentaires.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:{{region}}:{{account_id}}:secret:[[secret_id]]"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:{{region}}:{{account_id}}:key/[[key_id]]"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.*.amazonaws.com"
        ]
      }
    }
```

```
  },
  {
    "Effect": "Allow",
    "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
    ],
    "Resource": ["arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}",
 "arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}/data-source/*"]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}"
  }]
}
```

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM rôles pour les sources de données Jira

Lorsque vous utilisez Jira, vous attribuez un rôle avec les politiques suivantes.

- Autorisation d'accéder à votre AWS Secrets Manager secret pour authentifier votre Jira.
- Autorisation d'appeler le public requis APIs pour le connecteur Jira.
- Autorisation d'appeler le `BatchPutDocument`,`BatchDeleteDocument`,`PutPrincipalMapping`,`DeletePrincipalMapping`,`D` et `ListGroupsOlderThanOrderingId` APIs.

> **Note**
>
> Vous pouvez connecter une source de données Jira Amazon Kendra via Amazon VPC. Si vous utilisez un Amazon VPC, vous devez ajouter des autorisations supplémentaires.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:{{your-region}}:{{your-account-id}}:secret:
[[secret-id]]"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:{{your-region}}:{{your-account-id}}:key/[[key-id]]"
    ],
    "Condition": {
```

```
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.{{your-region}}.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
    ],
    "Resource": ["arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/
{{index-id}}", "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{index-
id}}/data-source/*"]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/
{{index-id}}"
  }]
}
```

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
```

```
        },
        "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM rôles pour les sources de données Microsoft Exchange

Lorsque vous utilisez une source de données Microsoft Exchange, vous fournissez Amazon Kendra un rôle doté des autorisations nécessaires pour se connecter au site. Il s'agit des licences suivantes :

- Autorisation d'obtenir et de déchiffrer le AWS Secrets Manager secret contenant l'ID d'application et la clé secrète nécessaires pour se connecter au site Microsoft Exchange. Pour plus d'informations sur le contenu du secret, consultez la section Sources de données Microsoft Exchange.
- Autorisation d'utiliser le BatchPutDocumentet BatchDeleteDocument APIs.

> **Note**
>
> Vous pouvez connecter une source de données Microsoft Exchange à Amazon Kendra via Amazon VPC. Si vous utilisez un Amazon VPC, vous devez ajouter des autorisations supplémentaires.

La IAM politique suivante fournit les autorisations nécessaires :

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:your-region:your-account-id:secret:secret-id"
```

```
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:your-region:your-account-id:key/key-id"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.your-region.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:your-region:your-account-id:index/index-id"
  }]
}
```

Si vous stockez la liste des utilisateurs à indexer dans un Amazon S3 compartiment, vous devez également autoriser l'utilisation de l'`GetObject`opération S3. La IAM politique suivante fournit les autorisations nécessaires :

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
```

```
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:your-region:your-account-id:secret:secret-id"
    ]
  },
  {
    "Action": [
      "s3:GetObject"
    ],
    "Resource": [
      "arn:aws:s3:::bucket-name/*"
    ],
    "Effect": "Allow"
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:your-region:your-account-id:key/[[key-ids]]"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.your-region.amazonaws.com",
          "s3.your-region.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:your-region:your-account-id:index/index-id"
  }]
}
```

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM rôles pour les sources OneDrive de données Microsoft

Lorsque vous utilisez une source de OneDrive données Microsoft, vous Amazon Kendra attribuez un rôle doté des autorisations nécessaires pour vous connecter au site. Il s'agit des licences suivantes :

- Autorisation d'obtenir et de déchiffrer le AWS Secrets Manager secret contenant l'ID de l'application et la clé secrète nécessaires pour se connecter au OneDrive site. Pour plus d'informations sur le contenu du secret, consultez la section Sources de OneDrive données Microsoft.
- Autorisation d'utiliser le BatchPutDocumentet BatchDeleteDocument APIs.

> **Note**
>
> Vous pouvez connecter une source de OneDrive données Microsoft à Amazon Kendra via Amazon VPC. Si vous utilisez un Amazon VPC, vous devez ajouter des autorisations supplémentaires.

La IAM politique suivante fournit les autorisations nécessaires :

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:your-region:your-account-id:secret:secret-id"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:your-region:your-account-id:key/key-id"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.your-region.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:your-region:your-account-id:index/index-id"
  }]
}
```

Si vous stockez la liste des utilisateurs à indexer dans un Amazon S3 compartiment, vous devez également autoriser l'utilisation de l'`GetObject`opération S3. La IAM politique suivante fournit les autorisations nécessaires :

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:your-region:your-account-id:secret:secret-id"
    ]
  },
  {
    "Action": [
      "s3:GetObject"
    ],
    "Resource": [
      "arn:aws:s3:::bucket-name/*"
    ],
    "Effect": "Allow"
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:your-region:your-account-id:key/[[key-ids]]"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.your-region.amazonaws.com",
          "s3.your-region.amazonaws.com"
        ]
      }
```

```
    }
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:your-region:your-account-id:index/index-id"
  }]
}
```

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM rôles pour les sources SharePoint de données Microsoft

IAM rôles pour SharePoint Connector v1.0

Pour une source de données Microsoft SharePoint Connector v1.0, vous fournissez un rôle avec les politiques suivantes.

- Autorisation d'accéder au AWS Secrets Manager secret qui contient le nom d'utilisateur et le mot de passe du SharePoint site. Pour plus d'informations sur le contenu du secret, consultez la section Sources de SharePoint données Microsoft.
- Autorisation d'utiliser la clé principale du AWS KMS client (CMK) pour déchiffrer le nom d'utilisateur et le secret du mot de passe stockés par. AWS Secrets Manager
- Autorisation d'utiliser les `BatchDeleteDocument` opérations `BatchPutDocument` et pour mettre à jour l'index.
- Autorisation d'accéder au Amazon S3 compartiment contenant le certificat SSL utilisé pour communiquer avec le SharePoint site.

Vous devez également joindre une politique de confiance qui permet Amazon Kendra d'assumer le rôle.

> **Note**
>
> Vous pouvez connecter une source de SharePoint données Microsoft à Amazon Kendra via Amazon VPC. Si vous utilisez un Amazon VPC, vous devez ajouter des autorisations supplémentaires.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "secretsmanager:GetSecretValue"
            ],
            "Resource": [
                "arn:aws:secretsmanager:your-region:your-account-
id:secret:secret-id"
            ]
        },
        {
            "Effect": "Allow",
            "Action": [
```

```
                "kms:Decrypt"
            ],
            "Resource": [
                "arn:aws:kms:your-region:your-account-id:key/key-id"
            ]
        },
        {
            "Effect": "Allow",
            "Action": [
                "kendra:BatchPutDocument",
                "kendra:BatchDeleteDocument"
            ],
            "Resource": [
                "arn:aws:kendra:your-region:your-account-id:index/index-id"
            ],
            "Condition": {
                "StringLike": {
                    "kms:ViaService": [
                        "kendra.your-region.amazonaws.com"
                    ]
                }
            }
        },
        {
            "Effect": "Allow",
            "Action": [
                "s3:GetObject"
            ],
            "Resource": [
                "arn:aws:s3:::bucket-name/*"
            ]
        }
    ]
}
```

Si vous avez chiffré le Amazon S3 compartiment contenant le certificat SSL utilisé pour communiquer avec le SharePoint site, définissez une politique pour donner Amazon Kendra accès à la clé.

JSON

```
{
```

```
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "kms:Decrypt"
            ],
            "Resource": [
                "arn:aws:kms:your-region:your-account-id:key/key-id"
            ]
        }
    ]
}
```

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM rôles pour SharePoint Connector v2.0

Pour une source de données Microsoft SharePoint Connector v2.0, vous fournissez un rôle avec les politiques suivantes.

- Autorisation d'accéder au AWS Secrets Manager secret qui contient les informations d'authentification du SharePoint site. Pour plus d'informations sur le contenu du secret, consultez la section Sources de SharePoint données Microsoft.

- Autorisation d'utiliser la clé principale du AWS KMS client (CMK) pour déchiffrer le nom d'utilisateur et le secret du mot de passe stockés par. AWS Secrets Manager
- Autorisation d'utiliser les `BatchDeleteDocument` opérations `BatchPutDocument` et pour mettre à jour l'index.
- Autorisation d'accéder au Amazon S3 compartiment contenant le certificat SSL utilisé pour communiquer avec le SharePoint site.

Vous devez également joindre une politique de confiance qui permet Amazon Kendra d'assumer le rôle.

> **Note**
>
> Vous pouvez connecter une source de SharePoint données Microsoft à Amazon Kendra via Amazon VPC. Si vous utilisez un Amazon VPC, vous devez ajouter des autorisations supplémentaires.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "secretsmanager:GetSecretValue"
      ],
      "Resource": [
        "arn:aws:secretsmanager:your-region:your-account-id:secret:secret-id"
      ]
    },
    {
      "Effect": "Allow",
      "Action": [
        "kms:Decrypt"
      ],
      "Resource": [
        "arn:aws:kms:your-region:your-account-id:key/key-id"
      ],
```

```
      "Condition": {
        "StringLike": {
          "kms:ViaService": [
            "secretsmanager.your-region.amazonaws.com"
          ]
        }
      }
    },
    {
      "Effect": "Allow",
      "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
      ],
      "Resource": [
        "arn:aws:kendra:your-region:your-account-id:index/index-id",
        "arn:aws:kendra:your-region:your-account-id:index/index-id/data-source/*"
      ]
    },
    {
      "Action": [
        "s3:GetObject"
      ],
      "Resource": [
        "arn:aws:s3:::bucket-name/key-name"
      ],
      "Effect": "Allow"
    },
    {
      "Effect": "Allow",
      "Action": [
        "kendra:BatchPutDocument",
        "kendra:BatchDeleteDocument"
      ],
      "Resource": "arn:aws:kendra:your-region:your-account-id:index/index-id"
    },
    {
      "Effect": "Allow",
      "Action": [
        "ec2:CreateNetworkInterface"
      ],
      "Resource": [
```

```
        "arn:aws:ec2:your-region:your-account-id:subnet/subnet-ids",
        "arn:aws:ec2:your-region:your-account-id:security-group/security-group"
      ]
    },
    {
      "Effect": "Allow",
      "Action": [
        "ec2:CreateNetworkInterface"
      ],
      "Resource": "arn:aws:ec2:region:account_id:network-interface/*",
      "Condition": {
        "StringLike": {
          "aws:RequestTag/AWS_KENDRA": "kendra_your-account-id_index-id_*"
        }
      }
    },
    {
      "Effect": "Allow",
      "Action": [
        "ec2:CreateTags"
      ],
      "Resource": "arn:aws:ec2:your-region:your-account-id:network-interface/*",
      "Condition": {
        "StringEquals": {
          "ec2:CreateAction": "CreateNetworkInterface"
        }
      }
    },

{
      "Effect": "Allow",
      "Action": [
        "ec2:CreateNetworkInterfacePermission"
      ],
      "Resource": "arn:aws:ec2:your-region:your-account-id:network-interface/*",
      "Condition": {
        "StringLike": {
          "aws:ResourceTag/AWS_KENDRA": "kendra_your-account-id_index-id_*"
        }
      }
    },
    {
      "Effect": "Allow",
      "Action": [
```

```
        "ec2:DescribeNetworkInterfaces",
        "ec2:DescribeAvailabilityZones",
        "ec2:DescribeNetworkInterfaceAttribute",
        "ec2:DescribeVpcs",
        "ec2:DescribeRegions",
        "ec2:DescribeNetworkInterfacePermissions",
        "ec2:DescribeSubnets"
      ],
      "Resource": "*"
    }
  ]
}
```

Si vous avez chiffré le Amazon S3 compartiment contenant le certificat SSL utilisé pour communiquer avec le SharePoint site, définissez une politique pour donner Amazon Kendra accès à la clé.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "kms:Decrypt"
            ],
            "Resource": [
                "arn:aws:kms:your-region:youraccount-id:key/key-id"
            ]
        }
    ]
}
```

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
```

```
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM rôles pour les sources de données Microsoft SQL Server

Lorsque vous utilisez Microsoft SQL Server, vous fournissez un rôle avec les politiques suivantes.

- Autorisation d'accéder à votre code AWS Secrets Manager secret pour authentifier votre instance Microsoft SQL Server.
- Autorisation d'appeler le public requis APIs pour le connecteur Microsoft SQL Server.
- Autorisation d'appeler le `BatchPutDocument`,`BatchDeleteDocument`,`PutPrincipalMapping`,`DeletePrincipalMapping`,`D` et `ListGroupsOlderThanOrderingId` APIs.

> **Note**
>
> Vous pouvez connecter une source de données Microsoft SQL Server à Amazon Kendra via Amazon VPC. Si vous utilisez un Amazon VPC, vous devez ajouter des autorisations supplémentaires.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
```

```
    ],
    "Resource": [
      "arn:aws:secretsmanager:{{region}}:{{account_id}}:secret:[[secret_id]]"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:{{region}}:{{account_id}}:key/[[key_id]]"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.*.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
    ],
    "Resource": ["arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}",
 "arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}/data-source/*"]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}"
  }]
}
```

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM rôles pour les sources de données Microsoft Teams

Lorsque vous utilisez une source de données Microsoft Teams, vous Amazon Kendra attribuez un rôle doté des autorisations nécessaires pour vous connecter au site. Il s'agit des licences suivantes :

- Autorisation d'obtenir et de déchiffrer le AWS Secrets Manager secret contenant l'ID client et le secret client nécessaires pour se connecter à Microsoft Teams. Pour plus d'informations sur le contenu du secret, consultez la section Sources de données Microsoft Teams.

> **Note**
>
> Vous pouvez connecter une source de données Microsoft Teams à Amazon Kendra via Amazon VPC. Si vous utilisez un Amazon VPC, vous devez ajouter des autorisations supplémentaires.

La IAM politique suivante fournit les autorisations nécessaires :

JSON

```
{
```

```
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:your-region:client-id:secret:secret-id"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:your-region:your-account-id:key/key-id"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.your-region.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:your-region:your-account-id:index/index-id"
  }]
}
```

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM rôles pour les sources de données Microsoft Yammer

Lorsque vous utilisez une source de données Microsoft Yammer, vous Amazon Kendra attribuez un rôle doté des autorisations nécessaires pour vous connecter au site. Il s'agit des licences suivantes :

- Autorisation d'obtenir et de déchiffrer le AWS Secrets Manager secret contenant l'ID de l'application et la clé secrète nécessaires pour se connecter au site Microsoft Yammer. Pour plus d'informations sur le contenu du secret, consultez la section Sources de données Microsoft Yammer.
- Autorisation d'utiliser le BatchPutDocumentet BatchDeleteDocument APIs.

> **Note**
>
> Vous pouvez connecter une source de données Microsoft Yammer à Amazon Kendra via Amazon VPC. Si vous utilisez un Amazon VPC, vous devez ajouter des autorisations supplémentaires.

La IAM politique suivante fournit les autorisations nécessaires :

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:your-region:your-account-id:secret:secret-id"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:your-region:your-account-id:key/key-id"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.your-region.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:your-region:your-account-id:index/index-id"
  }]
}
```

Si vous stockez la liste des utilisateurs à indexer dans un Amazon S3 compartiment, vous devez également autoriser l'utilisation de l'`GetObject`opération S3. La IAM politique suivante fournit les autorisations nécessaires :

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:your-region:your-account-id:secret:secret-id"
    ]
  },
  {
    "Action": [
      "s3:GetObject"
    ],
    "Resource": [
      "arn:aws:s3:::bucket-name/*"
    ],
    "Effect": "Allow"
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:your-region:your-account-id:key/[[key-ids]]"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.your-region.amazonaws.com",
          "s3.your-region.amazonaws.com"
        ]
      }
```

```
    }
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:your-region:your-account-id:index/index-id"
  }]
}
```

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM rôles pour les sources de données MySQL

Lorsque vous utilisez un connecteur de source de données My SQL, vous fournissez un rôle avec les politiques suivantes.

- Autorisation d'accéder à votre code AWS Secrets Manager secret pour authentifier votre instance de source de données My SQL.
- Autorisation d'appeler le public requis APIs pour le connecteur de source de données My SQL.

- Autorisation d'appeler le `BatchPutDocument`,`BatchDeleteDocument`,`PutPrincipalMapping`,`DeletePrincipalMapping`,`D` et `ListGroupsOlderThanOrderingId` APIs.

> **Note**
>
> Vous pouvez connecter une source de données MySQL à Amazon Kendra via Amazon VPC. Si vous utilisez un Amazon VPC, vous devez ajouter des autorisations supplémentaires.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:{{region}}:{{account_id}}:secret:[[secret_id]]"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:{{region}}:{{account_id}}:key/[[key_id]]"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.*.amazonaws.com"
        ]
      }
    }
  },
  {
```

```
    "Effect": "Allow",
    "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
    ],
    "Resource": ["arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}",
 "arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}/data-source/*"]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}"
  }]
}
```

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM rôles pour les sources de données Oracle

Lorsque vous utilisez un connecteur de source de données Oracle, vous fournissez un rôle avec les politiques suivantes.

- Autorisation d'accéder à votre code AWS Secrets Manager secret pour authentifier votre instance de source de données Oracle.
- Autorisation d'appeler le public requis APIs pour le connecteur de source de données Oracle.
- Autorisation d'appeler le `BatchPutDocument`,`BatchDeleteDocument`,`PutPrincipalMapping`,`DeletePrincipalMapping`,`D` et `ListGroupsOlderThanOrderingId` APIs.

> **Note**
>
> Vous pouvez connecter une source de données Oracle à Amazon Kendra via Amazon VPC. Si vous utilisez un Amazon VPC, vous devez ajouter des autorisations supplémentaires.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:{{region}}:{{account_id}}:secret:[[secret_id]]"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:{{region}}:{{account_id}}:key/[[key_id]]"
```

```
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.*.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
    ],
    "Resource": ["arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}",
 "arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}/data-source/*"]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}"
  }]
}
```

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
```

```
        },
        "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM rôles pour les sources de données PostgreSQL

Lorsque vous utilisez un connecteur de source de données PostgreSQL, vous fournissez un rôle avec les politiques suivantes.

- Autorisation d'accéder à votre code AWS Secrets Manager secret pour authentifier votre instance de source de données PostgreSQL.
- Autorisation d'appeler le public requis APIs pour le connecteur de source de données PostgreSQL.
- Autorisation d'appeler le `BatchPutDocument`,`BatchDeleteDocument`,`PutPrincipalMapping`,`DeletePrincipalMapping`,`D` et `ListGroupsOlderThanOrderingId` APIs.

> **Note**
>
> Vous pouvez connecter une source de données PostgreSQL à via. Amazon Kendra Amazon VPC Si vous utilisez un Amazon VPC, vous devez ajouter des autorisations supplémentaires.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:{{region}}:{{account_id}}:secret:[[secret_id]]"
    ]
  },
```

```
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:{{region}}:{{account_id}}:key/[[key_id]]"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.*.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
    ],
    "Resource": ["arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}",
 "arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}/data-source/*"]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}"
  }]
}
```

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM rôles pour les sources de données Quip

Lorsque vous utilisez Quip, vous attribuez un rôle avec les politiques suivantes.

- Autorisation d'accéder à votre AWS Secrets Manager secret pour authentifier votre Quip.
- Autorisation d'appeler le public requis APIs pour le connecteur Quip.
- Autorisation d'appeler le `BatchPutDocument`,`BatchDeleteDocument`,`PutPrincipalMapping`,`DeletePrincipalMapping`,`D` et `ListGroupsOlderThanOrderingId` APIs.

> **Note**
>
> Vous pouvez connecter une source de données Quip à Amazon Kendra via Amazon VPC. Si vous utilisez un Amazon VPC, vous devez ajouter des autorisations supplémentaires.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
```

```
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:{{your-region}}:{{your-account-id}}:secret:
[[secret-id]]"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:{{your-region}}:{{yoour-account-id}}:key/[[key-id]]"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.{{your-region}}.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
    ],
    "Resource": ["arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/
{{your-index-id}}", "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/
{{your-index-id}}/data-source/*"]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
```

```
    "Resource": "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/
{{index-id}}"
  }]
}
```

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM rôles pour les sources de données Salesforce

Lorsque vous utilisez Salesforce comme source de données, vous fournissez un rôle avec les politiques suivantes :

- Autorisation d'accéder au AWS Secrets Manager secret qui contient le nom d'utilisateur et le mot de passe du site Salesforce. Pour plus d'informations sur le contenu du secret, consultez la section Sources de données Salesforce.
- Autorisation d'utiliser la clé principale du AWS KMS client (CMK) pour déchiffrer le nom d'utilisateur et le secret du mot de passe stockés par. Secrets Manager
- Autorisation d'utiliser les `BatchDeleteDocument` opérations `BatchPutDocument` et pour mettre à jour l'index.

> **Note**
>
> Vous pouvez connecter une source de données Salesforce à Amazon Kendra via Amazon VPC. Si vous utilisez un Amazon VPC, vous devez ajouter des autorisations supplémentaires.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:your-region:your-account-id:secret:secret-id"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:your-region:your-account-id:key/key-id"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.your-region.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
```

```
    ],
    "Resource": "arn:aws:kendra:your-region:account-id:index/index-id"
  }]
}
```

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM rôles pour les sources ServiceNow de données

Lorsque vous utilisez un ServiceNow comme source de données, vous fournissez un rôle avec les politiques suivantes :

- Autorisation d'accéder au Secrets Manager secret qui contient le nom d'utilisateur et le mot de passe du ServiceNow site. Pour plus d'informations sur le contenu du secret, consultez la section Sources de ServiceNow données.
- Autorisation d'utiliser la clé principale du AWS KMS client (CMK) pour déchiffrer le nom d'utilisateur et le secret du mot de passe stockés par. Secrets Manager
- Autorisation d'utiliser les `BatchDeleteDocument` opérations `BatchPutDocument` et pour mettre à jour l'index.

> **Note**
>
> Vous pouvez connecter une source de ServiceNow données Amazon Kendra via Amazon VPC. Si vous utilisez un Amazon VPC, vous devez ajouter des autorisations supplémentaires.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:your-region:your-account-id:secret:secret-id"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:your-region:your-account-id:key/key-id"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.your-region.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
```

```
    ],
    "Resource": "arn:aws:kendra:your-region:your-account-id:index/index-id"
  }]
}
```

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM rôles pour les sources de données Slack

Lorsque vous utilisez Slack, vous attribuez un rôle soumis aux règles suivantes.

- Autorisation d'accéder à votre AWS Secrets Manager secret pour authentifier votre Slack.
- Autorisation d'appeler le public requis APIs pour le connecteur Slack.
- Autorisation d'appeler le `BatchPutDocument`,`BatchDeleteDocument`,`PutPrincipalMapping`,`DeletePrincipalMapping`,`D` et `ListGroupsOlderThanOrderingId` APIs.

> **Note**
>
> Vous pouvez connecter une source de données Slack à Amazon Kendra via Amazon VPC. Si vous utilisez un Amazon VPC, vous devez ajouter des autorisations supplémentaires.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:{{your-region}}:{{your-account-id}}:secret:
[[secret-id]]"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:{{your-region}}:{{account-id}}:key/[[key-id]]"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.{{region}}.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
    ],
    "Resource": ["arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/
{{index-id}}", "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{index-
id}}/data-source/*"]
  },
```

```
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/
{{index-id}}"
  }]
}
```

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM rôles pour les sources de données Zendesk

Lorsque vous utilisez Zendesk, vous attribuez un rôle avec les politiques suivantes.

- Autorisation d'accéder à votre AWS Secrets Manager code secret pour authentifier votre suite Zendesk.
- Autorisation d'appeler le public requis APIs pour le connecteur Zendesk.
- Autorisation d'appeler le `BatchPutDocument`,`BatchDeleteDocument`,`PutPrincipalMapping`,`DeletePrincipalMapping`,`D` et `ListGroupsOlderThanOrderingId` APIs.

> **Note**
>
> Vous pouvez connecter une source de données Zendesk à Amazon Kendra via Amazon VPC. Si vous utilisez un Amazon VPC, vous devez ajouter des autorisations supplémentaires.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:{{your-region}}:{{your-account-id}}:secret:
[[secret-id]]"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:{{your-region}}:{{your-account-id}}:key/[[key-id]]"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.{{your-region}}.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
        "kendra:PutPrincipalMapping",
```

```
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
    ],
    "Resource": ["arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/
{{index-id}}", "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{index-
id}}/data-source/*"]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/
{{index-id}}"
  }]
}
```

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

Rôle de cloud privé virtuel (VPC) IAM

Si vous utilisez un cloud privé virtuel (VPC) pour vous connecter à votre source de données, vous devez fournir les autorisations supplémentaires suivantes.

Rôle VPC IAM

```
{
      "Effect": "Allow",
      "Action": [
        "ec2:CreateNetworkInterface",
        "ec2:DeleteNetworkInterface"
      ],
      "Resource": [
        "arn:aws:ec2:{{region}}:{{account_id}}:subnet/[[subnet_ids]]",
        "arn:aws:ec2:{{region}}:{{account_id}}:security-group/[[security_group]]"
      ]
    },
    {
      "Effect": "Allow",
      "Action": [
        "ec2:CreateNetworkInterface",
        "ec2:DeleteNetworkInterface"
      ],
      "Resource": "arn:aws:ec2:{{region}}:{{account_id}}:network-interface/*",
      "Condition": {
        "StringLike": {
          "aws:RequestTag/AWS_KENDRA": "kendra_{{account_id}}_{{index_id}}_*"
        }
      }
    },
    {
      "Effect": "Allow",
      "Action": [
        "ec2:CreateTags"
      ],
      "Resource": "arn:aws:ec2:{{region}}:{{account_id}}:network-interface/*",
      "Condition": {
        "StringEquals": {
          "ec2:CreateAction": "CreateNetworkInterface"
        }
      }
    },
```

```
{
      "Effect": "Allow",
      "Action": [
        "ec2:CreateNetworkInterfacePermission"
      ],
      "Resource": "arn:aws:ec2:{{region}}:{{account_id}}:network-interface/*",
      "Condition": {
        "StringLike": {
          "aws:ResourceTag/AWS_KENDRA": "kendra_{{account_id}}_{{index_id}}_*"
        }
      }
    },
    {
      "Effect": "Allow",
      "Action": [
        "ec2:DescribeNetworkInterfaces",
        "ec2:DescribeAvailabilityZones",
        "ec2:DescribeNetworkInterfaceAttribute",
        "ec2:DescribeVpcs",
        "ec2:DescribeRegions",
        "ec2:DescribeNetworkInterfacePermissions",
        "ec2:DescribeSubnets"
      ],
      "Resource": "*"
    }
}
```

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
```

```
    ]
}
```

IAM rôles pour les questions fréquemment posées (FAQs)

Lorsque vous utilisez l'CreateFaqAPI pour charger des questions et réponses dans un index, vous devez fournir Amazon Kendra un IAM rôle ayant accès au Amazon S3 compartiment contenant les fichiers source. Si les fichiers source sont chiffrés, vous devez autoriser l'utilisation de la clé principale du AWS KMS client (CMK) pour déchiffrer les fichiers.

IAM rôles pour FAQs

Une politique de rôle obligatoire pour autoriser Amazon Kendra l'accès à un Amazon S3 compartiment.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "s3:GetObject"
            ],
            "Resource": [
                "arn:aws:s3:::bucket-name/*"
            ]
        }
    ]
}
```

Politique de rôle facultative autorisant l'utilisation Amazon Kendra d'une clé principale AWS KMS client (CMK) pour déchiffrer les fichiers d'un Amazon S3 compartiment.

JSON

```
{
```

```
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "kms:Decrypt"
            ],
            "Resource": [
                "arn:aws:kms:your-region:your-account-id:key/key-id"
            ],
            "Condition": {
                "StringLike": {
                    "kms:ViaService": [
                        "kendra.your-region.amazonaws.com"
                    ]
                }
            }
        }
    ]
}
```

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM rôles pour les suggestions de requêtes

Lorsque vous utilisez un Amazon S3 fichier comme liste de blocage de suggestions de requêtes, vous fournissez un rôle autorisé à accéder au Amazon S3 fichier et au Amazon S3 bucket. Si le fichier texte de la liste de blocage (le Amazon S3 fichier) du Amazon S3 compartiment est chiffré, vous devez autoriser l'utilisation de la clé principale du AWS KMS client (CMK) pour déchiffrer les documents.

IAM rôles pour les suggestions de requêtes

Une politique de rôle obligatoire Amazon Kendra pour autoriser l'utilisation du Amazon S3 fichier comme liste de blocage de suggestions de requêtes.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {"Effect": "Allow",
            "Action": [
                "s3:GetObject"
            ],
            "Resource": [
                "arn:aws:s3:::bucket-name/*"
            ]
        }
    ]
}
```

Politique de rôle facultative autorisant l'utilisation Amazon Kendra d'une clé principale AWS KMS du client (CMK) pour déchiffrer les documents d'un Amazon S3 compartiment.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {"Effect": "Allow",
```

```
            "Action": [
                "kms:Decrypt"
            ],
            "Resource": [
                "arn:aws:kms:your-region:your-account-id:key/key-id"
            ]
        }
    ]
}
```

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM rôles pour le mappage principal des utilisateurs et des groupes

Lorsque vous utilisez l'PutPrincipalMappingAPI pour associer des utilisateurs à leurs groupes afin de filtrer les résultats de recherche par contexte utilisateur, vous devez fournir une liste des utilisateurs ou des sous-groupes appartenant à un groupe. Si votre liste compte plus de 1 000 utilisateurs ou sous-groupes pour un groupe, vous devez fournir un rôle autorisé à accéder au Amazon S3 fichier de votre liste et au Amazon S3 bucket. Si le fichier texte (le Amazon S3 fichier) de la liste du Amazon S3 compartiment est chiffré, vous devez autoriser l'utilisation de la clé principale du AWS KMS client (CMK) pour déchiffrer les documents.

IAM rôles pour le mappage principal

Une politique de rôle obligatoire Amazon Kendra pour autoriser l'utilisation du Amazon S3 fichier comme liste d'utilisateurs et de sous-groupes appartenant à un groupe.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {"Effect": "Allow",
            "Action": [
                "s3:GetObject"
            ],
            "Resource": [
                "arn:aws:s3:::bucket-name/*"
            ]
        }
    ]
}
```

Politique de rôle facultative autorisant l'utilisation Amazon Kendra d'une clé principale AWS KMS du client (CMK) pour déchiffrer les documents d'un Amazon S3 compartiment.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {"Effect": "Allow",
            "Action": [
                "kms:Decrypt"
            ],
            "Resource": [
                "arn:aws:kms:your-region:your-account-id:key/key-id"
            ]
        }
    ]
}
```

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

Il est recommandé d'inclure `aws:sourceAccount` et `aws:sourceArn` dans la politique de confiance. Cela limite les autorisations et vérifie en toute sécurité si `aws:sourceAccount` et si `aws:sourceArn` elles sont identiques à celles prévues dans la politique de IAM rôle pour l'`sts:AssumeRole`action. Cela empêche les entités non autorisées d'accéder à vos IAM rôles et à leurs autorisations. Pour plus d'informations, consultez le AWS Identity and Access Management guide sur le problème de la confusion chez les députés.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Principal": {
                "Service": [
                    "kendra.amazonaws.com"
                ]
            },
```

```
            "Action": "sts:AssumeRole",
            "Condition": {
                "StringEquals": {
                    "aws:SourceAccount": "your-account-id"
                },
                "StringLike": {
                    "aws:SourceArn": "arn:aws:kendra:your-region:your-account-
id:index-id/*"
                }
            }
        }
    ]
}
```

IAM rôles pour AWS IAM Identity Center

Lorsque vous utilisez l'UserGroupResolutionConfigurationobjet pour récupérer les niveaux d'accès des groupes et des utilisateurs à partir d'une source d' AWS IAM Identity Center identité, vous devez fournir un rôle autorisé à y accéder IAM Identity Center.

IAM rôles pour AWS IAM Identity Center

Une politique de rôle obligatoire pour Amazon Kendra autoriser l'accès IAM Identity Center.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "sso-directory:SearchUsers",
                "sso-directory:ListGroupsForUser",
                "sso-directory:DescribeGroups",
                "sso:ListDirectoryAssociations"
            ],
            "Resource": [
                 "*"
            ]
        },
```

```
        {
          "Sid": "iamPassRole",
          "Effect": "Allow",
          "Action": "iam:PassRole",
          "Resource": "*",
          "Condition": {
            "StringEquals": {
              "iam:PassedToService": [
                "kendra.amazonaws.com"
              ]
            }
          }
        }
    ]
}
```

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM rôles liés aux Amazon Kendra expériences

Lorsque vous utilisez le CreateExperienceou UpdateExperience APIs pour créer ou mettre à jour une application de recherche, vous devez fournir un rôle autorisé à accéder aux opérations nécessaires et à IAM Identity Center.

IAM rôles liés à l'expérience Amazon Kendra de recherche

Une politique de rôle obligatoire pour autoriser l'accès Amazon Kendra aux `Query` opérations, aux `QuerySuggestions` opérations, aux `SubmitFeedback` opérations et au centre d'identité IAM qui stocke les informations de vos utilisateurs et de vos groupes.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "AllowsKendraSearchAppToCallKendraApi",
      "Effect": "Allow",
      "Action": [
        "kendra:GetQuerySuggestions",
        "kendra:Query",
        "kendra:DescribeIndex",
        "kendra:ListFaqs",
        "kendra:DescribeDataSource",
        "kendra:ListDataSources",
        "kendra:DescribeFaq",
        "kendra:SubmitFeedback"
      ],
      "Resource": [
        "arn:aws:kendra:your-region:your-account-id:index/index-id"
      ]
    },
    {
      "Sid": "AllowKendraSearchAppToDescribeDataSourcesAndFaq",
      "Effect": "Allow",
      "Action": [
        "kendra:DescribeDataSource",
        "kendra:DescribeFaq"
      ],
      "Resource": [
        "arn:aws:kendra:your-region:your-account-id:index/index-id/data-
source/data-source-id",
        "arn:aws:kendra:your-region:your-account-id:index/index-id/faq/faq-id"
      ]
    },
    {
```

```
      "Sid": "AllowKendraSearchAppToCallSSODescribeUsersAndGroups",
      "Effect": "Allow",
      "Action": [
        "sso-directory:ListGroupsForUser",
        "sso-directory:SearchGroups",
        "sso-directory:SearchUsers",
        "sso-directory:DescribeUser",
        "sso-directory:DescribeGroup",
        "sso-directory:DescribeGroups",
        "sso-directory:DescribeUsers",
        "sso:ListDirectoryAssociations"
      ],
      "Resource": [
        "*"
      ],
      "Condition": {
        "StringLike": {
          "kms:ViaService": [
            "kendra.your-region.amazonaws.com"
          ]
        }
      }
    }
  ]
}
```

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
```

```
}
```

Il est recommandé d'inclure `aws:sourceAccount` et `aws:sourceArn` dans la politique de confiance. Cela limite les autorisations et vérifie en toute sécurité si `aws:sourceAccount` et si `aws:sourceArn` elles sont identiques à celles prévues dans la politique de IAM rôle pour l'`sts:AssumeRole`action. Cela empêche les entités non autorisées d'accéder à vos IAM rôles et à leurs autorisations. Pour plus d'informations, consultez le AWS Identity and Access Management guide sur le problème de la confusion chez les députés.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Principal": {
                "Service": [
                    "kendra.amazonaws.com"
                ]
            },
            "Action": "sts:AssumeRole",
            "Condition": {
                "StringEquals": {
                    "aws:SourceAccount": "your-account-id"
                },
                "StringLike": {
                    "aws:SourceArn": "arn:aws:kendra:your-region:your-account-
id:index-id/*"
                }
            }
        }
    ]
}
```

IAM rôles pour l'enrichissement de documents personnalisés

Lorsque vous utilisez l'CustomDocumentEnrichmentConfigurationobjet pour appliquer des modifications avancées aux métadonnées et au contenu de votre

document, vous devez fournir un rôle doté des autorisations requises pour s'exécuter `PreExtractionHookConfiguration` et/ou`PostExtractionHookConfiguration`. Vous configurez une fonction Lambda pour `PreExtractionHookConfiguration` et/ou pour appliquer `PostExtractionHookConfiguration` des modifications avancées aux métadonnées et au contenu de votre document pendant le processus d'ingestion. Si vous choisissez d'activer le chiffrement côté serveur pour votre Amazon S3 compartiment, vous devez autoriser l'utilisation de la clé principale du AWS KMS client (CMK) pour chiffrer et déchiffrer les objets stockés dans votre compartiment. Amazon S3

IAM rôles pour l'enrichissement de documents personnalisés

Une politique de rôle obligatoire pour Amazon Kendra autoriser l'exécution `PreExtractionHookConfiguration` et `PostExtractionHookConfiguration` avec chiffrement pour votre Amazon S3 compartiment.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [{
    "Action": [
      "s3:GetObject",
      "s3:PutObject"
    ],
    "Resource": [
      "arn:aws:s3:::bucket-name/*"
    ],
    "Effect": "Allow"
  },
  {
    "Action": [
      "s3:ListBucket"
    ],
    "Resource": [
      "arn:aws:s3:::bucket-name"
    ],
    "Effect": "Allow"
  },
  {
    "Effect": "Allow",
    "Action": [
```

```
      "kms:Decrypt",
      "kms:GenerateDataKey"
    ],
    "Resource": [
      "arn:aws:kms:your-region:your-account-id:key/key-id"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "lambda:InvokeFunction"
    ],
    "Resource": "arn:aws:lambda:your-region:your-account-id:function:lambda-
function"
  }]
}
```

Une politique de rôle facultative autorisant Amazon Kendra l'exécution `PreExtractionHookConfiguration` `PostExtractionHookConfiguration` sans chiffrement de votre Amazon S3 compartiment.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [{
    "Action": [
      "s3:GetObject",
      "s3:PutObject"
    ],
    "Resource": [
      "arn:aws:s3:::bucket-name/*"
    ],
    "Effect": "Allow"
  },
  {
    "Action": [
      "s3:ListBucket"
    ],
    "Resource": [
      "arn:aws:s3:::bucket-name"
```

```
    ],
    "Effect": "Allow"
  },
  {
    "Effect": "Allow",
    "Action": [
      "lambda:InvokeFunction"
    ],
    "Resource": "arn:aws:lambda:your-region:your-account-id:function:lambda-
function"
  }]
}
```

Une politique de confiance pour Amazon Kendra permettre d'assumer un rôle.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

Il est recommandé d'inclure `aws:sourceAccount` et `aws:sourceArn` dans la politique de confiance. Cela limite les autorisations et vérifie en toute sécurité si `aws:sourceAccount` et si `aws:sourceArn` elles sont identiques à celles prévues dans la politique de IAM rôle pour l'`sts:AssumeRole`action. Cela empêche les entités non autorisées d'accéder à vos IAM rôles et à leurs autorisations. Pour plus d'informations, consultez le AWS Identity and Access Management guide sur le problème de la confusion chez les députés.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Principal": {
                "Service": [
                    "kendra.amazonaws.com"
                ]
            },
            "Action": "sts:AssumeRole",
            "Condition": {
                "StringEquals": {
                    "aws:SourceAccount": "your-account-id"
                },
                "StringLike": {
                    "aws:SourceArn": "arn:aws:kendra:your-region:your-account-
id:index-id/*"
                }
            }
        }
    ]
}
```

Déploiement Amazon Kendra

> **Note**
>
> La prise en charge des fonctionnalités varie en fonction du type d'index et de l'API de recherche utilisés. Pour savoir si cette fonctionnalité est prise en charge pour le type d'index et l'API de recherche que vous utilisez, consultez la section Types d'index.

Lorsque vient le temps de déployer Amazon Kendra la recherche sur votre site Web, nous fournissons le code source que vous pouvez utiliser avec React pour prendre une longueur d'avance sur votre application. Le code source est fourni gratuitement dans le cadre d'une licence MIT modifiée. Vous pouvez l'utiliser tel quel ou le modifier selon vos propres besoins. L'application React fournie est un exemple pour vous aider à démarrer. Il ne s'agit pas d'une application prête pour la production.

Pour déployer une application de recherche sans code et générer une URL de point de terminaison vers votre page de recherche avec contrôle d'accès, consultez Amazon Kendra Experience Builder.

L'exemple de code suivant ajoute une Amazon Kendra recherche à une application Web React existante :

- https://kendrasamples.s3.amazonaws.com/kendrasamples-react-app.zip —Exemples de fichiers que les développeurs peuvent utiliser pour créer une expérience de recherche fonctionnelle dans leur application Web React existante.

Les exemples sont calqués sur la page de recherche de la Amazon Kendra console. Ils ont les mêmes fonctionnalités de recherche et d'affichage des résultats de recherche. Vous pouvez utiliser l'exemple dans son intégralité ou choisir une seule des fonctionnalités pour votre usage personnel.

Pour afficher les trois composants de la page de recherche dans la Amazon Kendra console, choisissez l'icône du code (</>) dans le menu de droite. Passez le pointeur sur chaque section pour voir une brève description du composant et pour obtenir l'URL de la source du composant.

Rubriques

- Présentation
- Prérequis

- Configuration de l'exemple
- Page de recherche principale
- Composant de recherche
- Composante des résultats
- Composant Facettes
- Composant de pagination
- Création d'une expérience de recherche sans code

Présentation

Vous ajoutez l'exemple de code à une application Web React existante pour activer la recherche. L'exemple de code inclut un fichier Readme avec les étapes à suivre pour configurer un nouvel environnement de développement React. Les données d'exemple contenues dans l'exemple de code peuvent être utilisées pour illustrer une recherche. Les fichiers et composants de recherche de l'exemple de code sont structurés comme suit :

- Page de recherche principale (`Search.tsx`) : il s'agit de la page principale qui contient tous les composants. C'est ici que vous intégrez votre application à l' Amazon Kendra API.
- Barre de recherche : il s'agit du composant dans lequel un utilisateur saisit un terme de recherche et appelle la fonction de recherche.
- Résultats : il s'agit du composant qui affiche les résultats de. Amazon Kendra Il comporte trois éléments : les réponses suggérées, les résultats des FAQ et les documents recommandés.
- Facettes : il s'agit du composant qui affiche les facettes dans les résultats de recherche et vous permet de choisir une facette pour affiner la recherche.
- Pagination : il s'agit du composant à partir duquel la réponse est paginée. Amazon Kendra

Prérequis

Avant de commencer, vous avez besoin de ce qui suit :

- Node.js et npm installés. La version 19 ou antérieure de Node.js est requise.
- Python 3 ou Python 2 téléchargé et installé.
- SDK pour Javaou AWS SDK pour JavaScriptpour effectuer des appels d'API à Amazon Kendra.

- Une application Web React existante. L'exemple de code inclut un fichier Readme avec les étapes à suivre pour configurer un nouvel environnement de développement React, notamment en utilisant les frameworks/bibliothèques requis. Vous pouvez également suivre les instructions de démarrage rapide de la documentation React sur la création d'une application Web React.
- Les bibliothèques et dépendances requises sont configurées dans votre environnement de développement. L'exemple de code inclut un fichier Readme qui répertorie les bibliothèques requises et les dépendances des packages. Notez que cela `sass` est obligatoire, car cela `node-sass` est obsolète. Si vous l'avez déjà installé`node-sass`, désinstallez-le et installez-le`sass`.

Configuration de l'exemple

Une procédure complète pour ajouter une Amazon Kendra recherche à une application React se trouve dans le fichier Readme inclus dans l'exemple de code.

Pour commencer à utiliser le kendrasamples-react-app fichier .zip

1. Assurez-vous d'avoir terminé lePrérequis, y compris le téléchargement et l'installation de Node.js et de npm.
2. Téléchargez le kendrasamples-react-app fichier .zip et décompressez-le.
3. Ouvrez votre terminal et accédez à`aws-kendra-example-react-app/src/services/`. Ouvrez `local-dev-credentials.json` et saisissez vos informations d'identification. N'ajoutez ce fichier à aucun dépôt public.
4. Accédez à `aws-kendra-example-react-app` et installez les dépendances dans`package.json`. Exécutez `npm install`.
5. Lancez une version de démonstration de votre application sur votre serveur local. Exécutez `npm start`. Vous pouvez arrêter le serveur local en tapant sur votre clavier`Cmd/Ctrl + C`.
6. Vous pouvez modifier le port ou l'hôte (par exemple, l'adresse IP) en accédant à `package.json` et en mettant à jour l'hôte et le port :`"start": "HOST=[host] PORT=[port] react-scripts start"`. Si vous utilisez Windows :`"start": "set HOST=[host] && set PORT=[port] && react-scripts start"`.
7. Si vous avez un domaine de site Web enregistré, vous pouvez le spécifier `package.json` après le nom de votre application. Par exemple, `"homepage": "https://mywebsite.com"`. Vous devez exécuter à `npm install` nouveau pour mettre à jour les nouvelles dépendances, puis exécuter`npm start`.

8. Pour créer l'application, exécutez`npm build`. Téléchargez le contenu du répertoire de construction sur votre fournisseur d'hébergement.

> **Warning**
>
> L'application React n'est pas prête pour la production. Il s'agit d'un exemple de déploiement d'une application à des fins Amazon Kendra de recherche.

Page de recherche principale

La page de recherche principale (`Search.tsx`) contient tous les exemples de composants de recherche. Il inclut le composant de barre de recherche pour la sortie, les composants de résultats pour afficher la réponse de l'API Query et un composant de pagination pour parcourir la réponse.

Composant de recherche

Le composant de recherche fournit une zone de texte pour saisir le texte de la requête. La `onSearch` fonction est un hook qui appelle la fonction principale `Search.tsx` pour effectuer l'appel de l'API Amazon Kendra Query.

Composante des résultats

Le composant de résultats affiche la réponse de l'`Query`API. Les résultats sont présentés dans trois zones distinctes.

- Réponses suggérées : il s'agit des meilleurs résultats renvoyés par l'`Query`API. Il contient jusqu'à trois suggestions de réponses. Dans la réponse, ils ont le type de résultat`ANSWER`.
- Réponses aux FAQ : il s'agit des réponses aux questions fréquemment posées renvoyées par la réponse. FAQs sont ajoutés à l'index séparément. Dans la réponse, ils ont le type`QUESTION_ANSWER`. Pour plus d'informations, consultez la section Questions et réponses.
- Documents recommandés : il s'agit de documents supplémentaires Amazon Kendra renvoyés dans la réponse. Dans la réponse de l'`Query`API, ils ont le type`DOCUMENT`.

Les composants des résultats partagent un ensemble de composants pour des fonctionnalités telles que le surlignage, les titres, les liens, etc. Les composants partagés doivent être présents pour que les composants du résultat fonctionnent.

Composant Facettes

Le composant facettes répertorie les facettes disponibles dans les résultats de recherche. Chaque facette classe la réponse selon une dimension spécifique, telle que l'auteur. Vous pouvez affiner la recherche sur une facette spécifique en en choisissant une option dans la liste.

Une fois que vous avez sélectionné une facette, le composant appelle `Query` avec un filtre d'attributs qui limite la recherche aux documents correspondant à cette facette.

Composant de pagination

Le composant de pagination vous permet d'afficher les résultats de recherche de l'`Query`API sur plusieurs pages. Il appelle l'`Query`API avec les `PageNumber` paramètres `PageSize` et pour obtenir une page de résultats spécifique.

Création d'une expérience de recherche sans code

Vous pouvez créer et déployer une application Amazon Kendra de recherche sans avoir besoin de code frontal. Amazon Kendra Experience Builder vous aide à créer et à déployer une application de recherche entièrement fonctionnelle en quelques clics afin que vous puissiez commencer à rechercher immédiatement. Vous pouvez personnaliser votre page de recherche et ajuster votre recherche pour adapter l'expérience aux besoins de vos utilisateurs. Amazon Kendra génère une URL de point de terminaison unique et entièrement hébergée de votre page de recherche pour commencer à rechercher vos documents et FAQs. Vous pouvez rapidement établir une preuve de concept de votre expérience de recherche et la partager avec d'autres personnes.

Vous utilisez le modèle d'expérience de recherche disponible dans le générateur pour personnaliser votre recherche. Vous pouvez inviter d'autres personnes à collaborer au développement de votre expérience de recherche ou à évaluer les résultats de recherche à des fins d'optimisation. Une fois que votre expérience de recherche est prête pour que vos utilisateurs puissent commencer à effectuer des recherches, il vous suffit de partager l'URL sécurisée du point de terminaison.

Fonctionnement du moteur de recherche Experience Builder

Le processus global de création d'une expérience de recherche est le suivant :

1. Vous créez votre expérience de recherche en lui attribuant un nom, une description et en choisissant les sources de données que vous souhaitez utiliser pour votre expérience de recherche.

2. Vous configurez votre liste d'utilisateurs et de groupes, AWS IAM Identity Center puis vous leur attribuez des droits d'accès à votre expérience de recherche. Vous vous incluez en tant que propriétaire de l'expérience. Pour de plus amples informations, veuillez consulter the section called “Fournir un accès à votre page de recherche”.
3. Vous ouvrez l' Amazon Kendra Experience Builder pour concevoir et ajuster votre page de recherche. Vous pouvez partager l'URL de votre point de terminaison correspondant à votre expérience de recherche avec d'autres personnes à qui vous attribuez des droits d'accès de modification ou des droits d'accès de consultation et de recherche.

Vous appelez l'CreateExperienceAPI pour créer et configurer votre expérience de recherche. Si vous utilisez la console, vous sélectionnez votre index, puis Experiences dans le menu de navigation pour configurer votre expérience.

Concevez et optimisez votre expérience de recherche

Une fois que vous avez créé et configuré votre expérience de recherche, vous l'ouvrez à l'aide d'une URL de point de terminaison pour commencer à personnaliser votre recherche en tant que propriétaire disposant de droits d'accès aux éditeurs. Vous tapez votre requête dans le champ de recherche, puis vous personnalisez votre recherche à l'aide des options d'édition du panneau latéral pour voir comment elles s'appliquent à votre page. Lorsque vous êtes prêt à publier, sélectionnez Publier. Vous pouvez également basculer entre le mode Live View, pour afficher la dernière version publiée de votre page de recherche, et le Basculer en mode build, pour modifier ou personnaliser votre page de recherche.

Les méthodes suivantes vous permettent de personnaliser votre expérience de recherche.

Filtre

Ajoutez une recherche à facettes ou filtrez par attributs de document. Cela inclut les attributs personnalisés. Vous pouvez ajouter un filtre à l'aide de vos propres champs de métadonnées configurés. Par exemple, pour effectuer une recherche par facettes par catégorie de ville, utilisez un attribut de document `_category` personnalisé qui contient toutes les catégories de villes.

Réponse suggérée

Ajoutez des réponses générées par le machine learning aux requêtes de vos utilisateurs. Par exemple, « Quelle est la difficulté de ce cours ? » . Amazon Kendra peut récupérer le texte le plus pertinent parmi tous les documents faisant référence à la difficulté d'un cours et suggérer la réponse la plus pertinente.

FAQ

Ajoutez un document FAQ pour fournir des réponses aux questions fréquemment posées. Par exemple, « Combien d'heures faudra-t-il pour suivre ce cours ? » . Amazon Kendra peut utiliser le document FAQ contenant la réponse à cette question et donner la bonne réponse.

Tri

Ajoutez le tri des résultats de recherche afin que vos utilisateurs puissent organiser les résultats par pertinence, date de création, date de dernière mise à jour et autres critères de tri.

Documents

Configurez le mode d'affichage des documents ou des résultats de recherche sur votre page de recherche. Vous pouvez configurer le nombre de résultats affichés sur la page, inclure une pagination telle que les numéros de page, activer un bouton de commentaires utilisateur et organiser la manière dont les champs de métadonnées des documents sont affichés dans un résultat de recherche.

Langue

Sélectionnez une langue pour filtrer les résultats de recherche ou les documents dans la langue sélectionnée.

Boîte de recherche

Configurez la taille et le texte de remplacement de votre champ de recherche, et autorisez les suggestions de requêtes.

Réglage de la pertinence

Renforcez les champs de métadonnées des documents pour donner plus de poids à ces champs lorsque vos utilisateurs recherchent des documents. Vous pouvez ajouter un poids qui commence à 1 et augmente progressivement jusqu'à 10. Vous pouvez améliorer les types de champs texte, de date et numériques. Par exemple, pour donner `_created_at` plus `_last_updated_at` de poids ou d'importance qu'aux autres champs, attribuez-leur une pondération de 1 à 10, en fonction de leur importance. Vous pouvez appliquer différentes configurations de réglage de la pertinence pour chaque application ou expérience de recherche.

Fournir un accès à votre page de recherche

L'accès à votre expérience de recherche se fait via IAM Identity Center. Lorsque vous configurez votre expérience de recherche, vous autorisez les autres personnes répertoriées dans votre répertoire Identity Center à accéder à votre page Amazon Kendra de recherche. Ils reçoivent un e-mail leur demandant de se connecter à l'aide de leurs informations d'identification dans IAM Identity Center pour accéder à la page de recherche. Vous devez configurer IAM Identity Center au niveau de l'organisation ou du titulaire du compte dans AWS Organizations. Pour plus d'informations sur la configuration d'IAM Identity Center, consultez Getting started with IAM Identity Center.

Vous activez les identités des utilisateurs dans IAM Identity Center en fonction de votre expérience de recherche et vous attribuez des autorisations d'accès au Viewer ou au Owner à l'aide de l'API ou de la console.

- Afficheur : autorisé à émettre des requêtes, à recevoir des suggestions de réponses pertinentes pour sa recherche et à apporter ses commentaires Amazon Kendra afin d'améliorer continuellement la recherche.
- Propriétaire : autorisé à personnaliser le design de la page de recherche, à ajuster la recherche et à utiliser l'application de recherche en tant que visualiseur. La désactivation de l'accès aux spectateurs dans la console n'est actuellement pas prise en charge.

Pour attribuer à d'autres personnes l'accès à votre expérience de recherche, vous devez d'abord activer les identités des utilisateurs dans IAM Identity Center en fonction de votre Amazon Kendra expérience en utilisant l'ExperienceConfigurationobjet. Vous spécifiez le nom du champ qui contient les identifiants de vos utilisateurs, tels que le nom d'utilisateur ou l'adresse e-mail. Vous autorisez ensuite votre liste d'utilisateurs à accéder à votre expérience de recherche à l'aide de l'AssociateEntitiesToExperienceAPI et définissez leurs autorisations en tant que Viewer ou Owner à l'aide de l'AssociatePersonasToEntitiesAPI. Vous spécifiez chaque utilisateur ou groupe utilisant l'EntityConfigurationobjet et indiquez si cet utilisateur ou ce groupe est un visualiseur ou un propriétaire utilisant l'EntityPersonaConfiguratonobjet.

Pour autoriser d'autres personnes à accéder à votre expérience de recherche à l'aide de la console, vous devez d'abord créer une expérience, confirmer votre identité et confirmer que vous êtes propriétaire. Vous pouvez ensuite désigner d'autres utilisateurs ou groupes en tant que spectateurs ou propriétaires. Dans la console, sélectionnez votre index, puis sélectionnez Expériences dans le menu de navigation. Après avoir créé votre expérience, vous pouvez la sélectionner dans la liste.

Accédez à Gestion des accès pour attribuer des utilisateurs ou des groupes en tant que spectateurs ou propriétaires.

Configuration d'une expérience de recherche

Voici un exemple de configuration ou de création d'une expérience de recherche.

Console

Pour créer une expérience Amazon Kendra de recherche

1. Dans le volet de navigation de gauche, sous Indexes, sélectionnez Expériences, puis sélectionnez Créer une expérience.
2. Sur la page Configurer l'expérience, entrez le nom et la description de votre expérience, choisissez vos sources de contenu et choisissez le rôle IAM pour votre expérience. Pour plus d'informations sur les rôles IAM, consultez la section Rôles IAM pour Amazon Kendra les expériences.
3. Sur la page Confirmez votre identité à partir d'un annuaire Identity Center, sélectionnez votre nom d'utilisateur tel que votre adresse e-mail. Si vous n'avez pas de répertoire Identity Center, entrez simplement votre nom complet et votre adresse e-mail pour créer un annuaire Identity Center. Cela vous inclut en tant qu'utilisateur de l'expérience et vous attribue automatiquement les droits d'accès de propriétaire.
4. Sur la page Vérifier pour ouvrir Experience Builder, passez en revue les détails de votre configuration, sélectionnez Créer une expérience et ouvrez Experience Builder pour commencer à modifier votre page de recherche.

CLI

Pour créer une Amazon Kendra expérience

```
aws kendra create-experience \
 --name experience-name \
 --description "experience description" \
 --index-id index-id \
 --role-arn arn:aws:iam::account-id:role/role-name \
 --configuration '{"ExperienceConfiguration":[{"ContentSourceConfiguration":
{"DataSourceIds":["data-source-1","data-source-2"]},
 "UserIdentityConfiguration":"identity attribute name"}]}'
```

```
aws kendra describe-experience \
 --endpoints experience-endpoint-URL(s)
```

Python

Pour créer une Amazon Kendra expérience

```
import boto3
from botocore.exceptions import ClientError
import pprint
import time

kendra = boto3.client("kendra")

print("Create an experience.")

# Provide a name for the experience
name = "experience-name"
# Provide an optional description for the experience
description = "experience description"
# Provide the index ID for the experience
index_id = "index-id"
# Provide the IAM role ARN required for Amazon Kendra experiences
role_arn = "arn:aws:iam::${account-id}:role/${role-name}"
# Configure the experience
configuration = {"ExperienceConfiguration":
        [{
            "ContentSourceConfiguration":{"DataSourceIds":["data-source-1","data-
source-2"]},
            "UserIdentityConfiguration":"identity attribute name"
        }]
    }

try:
    experience_response = kendra.create_experience(
        Name = name,
        Description = description,
        IndexId = index_id,
        RoleArn = role_arn,
        Configuration = configuration
    )

    pprint.pprint(experience_response)
```

```
    experience_endpoints = experience_response["Endpoints"]

    print("Wait for Amazon Kendra to create the experience.")

    while True:
        # Get the details of the experience, such as the status
        experience_description = kendra.describe_experience(
            Endpoints = experience_endpoints
        )
        status = experience_description["Status"]
        print(" Creating experience. Status: "+status)
        time.sleep(60)
        if status != "CREATING":
            break

except  ClientError as e:
        print("%s" % e)

print("Program ends.")
```

Java

Pour créer un Amazon Kendra

```
package com.amazonaws.kendra;

import java.util.concurrent.TimeUnit;
import software.amazon.awssdk.services.kendra.KendraClient;
import software.amazon.awssdk.services.kendra.model.CreateExperienceRequest;
import software.amazon.awssdk.services.kendra.model.CreateExperienceResponse;
import software.amazon.awssdk.services.kendra.model.DescribeExperienceRequest;
import software.amazon.awssdk.services.kendra.model.DescribeExperienceResponse;
import software.amazon.awssdk.services.kendra.model.ExperienceStatus;


public class CreateExperienceExample {

    public static void main(String[] args) throws InterruptedException {
        System.out.println("Create an experience");

        String experienceName = "experience-name";
        String experienceDescription = "experience description";
        String indexId = "index-id";
        String experienceRoleArn = "arn:aws:iam::account-id:role/role-name";
```

```
        KendraClient kendra = KendraClient.builder().build();

        CreateExperienceRequest createExperienceRequest = CreateExperienceRequest
            .builder()
            .name(experienceName)
            .description(experienceDescription)
            .roleArn(experienceRoleArn)
            .configuration(
                ExperienceConfiguration
                    .builder()
                    .contentSourceConfiguration(
                        ContentSourceConfiguration(
                            .builder()
                            .dataSourceIds("data-source-1","data-source-2")
                            .build()
                        )
                    )
                    .userIdentityConfiguration(
                        UserIdentityConfiguration(
                            .builder()
                            .identityAttributeName("identity-attribute-name")
                            .build()
                        )
                    ).build()
            ).build();

        CreateExperienceResponse createExperienceResponse =
kendra.createExperience(createExperienceRequest);
        System.out.println(String.format("Experience response %s",
createExperienceResponse));

        String experienceEndpoints = createExperienceResponse.endpoints();

        System.out.println(String.format("Wait for Kendra to create the
experience.", experienceEndpoints));
        while (true) {
            DescribeExperienceRequest describeExperienceRequest =
DescribeExperienceRequest.builder().endpoints(experienceEndpoints).build();
            DescribeExperienceResponse describeEpxerienceResponse =
kendra.describeExperience(describeExperienceRequest);
            ExperienceStatus status = describeExperienceResponse.status();
            TimeUnit.SECONDS.sleep(60);
            if (status != ExperienceStatus.CREATING) {
```

```
                break;
            }
        }

        System.out.println("Experience creation is complete.");
    }
}
```

Ajustement de la capacité

> **Note**
>
> La prise en charge des fonctionnalités varie en fonction du type d'index et de l'API de recherche utilisés. Pour savoir si cette fonctionnalité est prise en charge pour le type d'index et l'API de recherche que vous utilisez, consultez la section Types d'index.

Amazon Kendra fournit des ressources pour votre index en unités de capacité. Chaque unité de capacité fournit des ressources supplémentaires pour votre index. Il existe des unités de capacité distinctes pour le stockage des documents et pour les requêtes. Vous ne pouvez ajouter des unités de capacité qu'aux indices GenAI Enterprise Edition et Amazon Kendra Enterprise. Vous ne pouvez pas ajouter de capacité à un index Developer Edition.

Une unité de capacité de stockage de documents fournit le stockage supplémentaire suivant pour votre index.

- Amazon Kendra GenAI Enterprise Edition — 20 000 documents ou 200 Mo de texte extrait.
- Amazon Kendra Enterprise Edition : 100 000 documents ou 30 Go de stockage.

Une unité de capacité de requête fournit les requêtes supplémentaires suivantes pour votre index. Les requêtes par seconde sont partagées entre les API de récupération et de requête.

- Amazon Kendra GenAI Enterprise Edition : 0,1 requête par seconde, soit environ 8 000 requêtes par jour.
- Amazon Kendra Enterprise Edition : 0,1 requête par seconde, soit environ 8 000 requêtes par jour.

Chaque index est doté d'une capacité de base égale à 1 unité de capacité (30 Go/200Mo de stockage et 0,1 requête par seconde). Il y a un coût supplémentaire pour chaque unité de capacité supplémentaire. Pour de plus amples informations, veuillez consulter Tarification Amazon Kendra.

Vous pouvez ajouter jusqu'à 100 unités de capacité supplémentaires à votre espace de stockage et interroger des ressources pour obtenir un index. Si vous avez besoin de plus d'unités, contactez simplement le Support.

Vous pouvez ajuster les unités de capacité jusqu'à 5 fois par jour pour répondre à vos besoins d'utilisation. Vous ne pouvez pas réduire la capacité de stockage des documents en dessous du nombre de documents stockés dans votre index. Par exemple, si vous stockez 150 000 documents, vous ne pouvez pas réduire la capacité de stockage en dessous d'une unité supplémentaire.

Vous pouvez afficher les ressources utilisées par un index dans la console en sélectionnant le nom de l'index pour ouvrir les paramètres de l'index et d'autres informations, ou vous pouvez utiliser l'DescribeIndexAPI.

Amazon Kendra renvoie également des exceptions lorsque vous dépassez la capacité d'un index. Vous obtenez un `ServiceQuotaExceededException` lorsque la taille totale extraite de tous les documents dépasse la limite d'un index. Vous obtenez un `InvalidRequest` pour chaque document lorsque le nombre de documents dépasse la limite d'un index. Vous obtenez un `ThrottlingException` lorsque le nombre de requêtes par seconde dépasse la limite. Pour plus d'informations sur les limites, consultez la section Quotas pour Amazon Kendra.

Les requêtes accumulées dureront jusqu'à 24 heures.

Capacité de visionnage

Consultez les ressources utilisées par votre index à l'aide de la Amazon Kendra console en sélectionnant le nom de votre index pour accéder aux détails. La console fournit également des graphiques d'utilisation qui vous permettent de déterminer la capacité de stockage et de requête utilisée par votre index. Vous pouvez utiliser ces informations pour vous aider à planifier à quel moment ajouter de la capacité supplémentaire.

Pour afficher le stockage des documents et utiliser les requêtes (console)

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la à la https://console.aws.amazon.com/kendra/maison.
2. Dans la liste des index, choisissez l'index auquel vous souhaitez accéder.
3. Accédez à la section des paramètres pour afficher le stockage total actuel des documents et la capacité de requête.

Pour afficher la capacité à l'aide de l' Amazon Kendra API, utilisez le `CapacityUnits` paramètre de l'DescribeIndexAPI.

Ajouter et supprimer de la capacité

Si vous avez besoin d'une capacité supplémentaire pour votre index, vous pouvez l'ajouter à l'aide de la console ou de l' Amazon Kendra API.

Pour ajouter ou supprimer de la capacité de stockage ou de requête (console)

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la à la https://console.aws.amazon.com/kendra/maison.
2. Dans la liste des index, choisissez l'index auquel vous souhaitez accéder.
3. Sélectionnez Modifier ou sélectionnez Modifier dans le menu déroulant Actions.
4. Sélectionnez Suivant pour accéder à la page des détails du provisionnement.
5. Ajoutez ou supprimez des unités de capacité de stockage de documents et/ou de capacité de requête.
6. Continuez à sélectionner Suivant pour accéder à la page de révision, puis sélectionnez Mettre à jour pour enregistrer vos modifications.

Une fois que vous avez mis à jour la capacité de votre index, plusieurs minutes peuvent être nécessaires pour que les modifications prennent effet.

Pour ajouter ou supprimer de la capacité à l'aide de l' Amazon Kendra API, utilisez le `CapacityUnits` paramètre de l'UpdateIndexAPI.

Amazon Kendra Capacité de classement intelligente

Une unité de capacité fournit les demandes de renotation supplémentaires suivantes par seconde pour un plan d'exécution de la renotation. Un plan d'exécution Rescore est une ressource utilisée pour approvisionner l'API Rescore.

- 0,01 requêtes par seconde.

Chaque plan d'exécution du rescore est doté d'une capacité de base égale à 1 unité de capacité (0,01 requêtes par seconde). Il y a un coût supplémentaire pour chaque unité de capacité supplémentaire. Pour de plus amples informations, veuillez consulter Tarification Amazon Kendra.

Vous pouvez ajouter jusqu'à 1 000 unités de capacité supplémentaires pour un plan d'exécution de la nouvelle notation. Si vous avez besoin de plus d'unités, contactez simplement le Support.

Capacité de suggestions de requêtes

Lorsque vous utilisez des suggestions de requêtes, la capacité de requête de base est de 2,5 GetQuerySuggestionsappels par seconde. La `GetQuerySuggestions` capacité est cinq fois supérieure à la capacité de requête allouée pour un index, ou à la capacité de base de 2,5 appels par seconde, selon la valeur la plus élevée. Par exemple, la capacité de base d'un index est de 0,1 requête par seconde, et la capacité `GetQuerySuggestions` a une base de 2,5 appels par seconde. Si vous ajoutez 0,1 requête supplémentaire par seconde pour un total de 0,2 requête par seconde pour un index, la capacité `GetQuerySuggestions` est de 2,5 appels par seconde (supérieure à cinq fois 0,2 requête par seconde).

Amazon Kendra capacité d'expérience

Capacité d'expérience de recherche

Amazon Kendra commence à s'accélérer `QueryQuerySuggestions`, `SubmitFeedback` pour votre Amazon Kendra expérience à 15 requêtes par seconde et à 40 demandes par seconde pour l'éclatement des requêtes. Pour un index comportant plus de 150 unités de capacité de requête, ces limites s'appliquent toujours.

Par exemple, vos unités de capacité de requête pour votre index sont de 150, de sorte que votre application d'expérience de recherche peut traiter 15 requêtes par seconde. Toutefois, si vous passiez à 200 unités de capacité de requête, votre application d'expérience de recherche ne traiterait toujours que 15 requêtes par seconde. Si vous limitez votre index à 100 unités de capacité de requête, votre application d'expérience de recherche ne traitera que 10 requêtes par seconde.

rafale de requêtes adaptative

Amazon Kendra possède une capacité de base provisionnée de 1 unité de capacité de requête. Vous pouvez utiliser jusqu'à 8 000 requêtes par jour avec un débit minimum de 0,1 requête par seconde (par unité de capacité de requête). Les requêtes accumulées dureront jusqu'à 24 heures et peuvent faire face à des pics de trafic. La quantité de rafale autorisée varie car elle dépend de la charge du cluster à un moment donné. Fournissez suffisamment d'unités de capacité de requête pour gérer vos pics de charge.

Une approche adaptative permettant de gérer les pics de trafic inattendus au-delà du débit fourni est la mise en rafale Amazon Kendra de requêtes adaptative intégrée. La fonction Adaptive Query Bursting est disponible dans l'édition Enterprise de Amazon Kendra.

Le rafistage de requêtes adaptatif est une fonctionnalité intégrée qui vous permet d'appliquer une capacité de requête inutilisée pour gérer le trafic inattendu. Amazon Kendra accumule vos requêtes inutilisées au rythme des requêtes provisionnées par seconde, chaque seconde, jusqu'au nombre maximum de requêtes que vous avez provisionnées pour votre index. Amazon Kendra Ces requêtes accumulées sont utilisées pour le trafic inattendu supérieur à la capacité allouée. Les performances optimales de la compression adaptative des requêtes peuvent varier en fonction de plusieurs facteurs tels que la taille totale de votre index, la complexité des requêtes, le cumul de requêtes inutilisées et la charge globale de votre index. Il est recommandé d'effectuer vos propres tests de charge pour mesurer avec précision la capacité d'éclatement.

Premiers pas

Cette section explique comment créer une source de données et ajouter vos documents à un Amazon Kendra index. Des instructions sont fournies pour la AWS console AWS CLI, un programme Python utilisant le AWS SDK pour Python (Boto3), et un programme Java utilisant le AWS SDK pour Java.

Rubriques

- Prérequis
- Commencer à utiliser la Amazon Kendra console
- Démarrer (AWS CLI)
- Démarrer (AWS SDK pour Python (Boto3))
- Démarrer (AWS SDK pour Java)
- Commencer à utiliser une source de Amazon S3 données (console)
- Commencer à utiliser une source de données de base de données MySQL (console)
- Commencer à utiliser une source AWS IAM Identity Center d'identité (console)

Prérequis

Les étapes suivantes sont des prérequis pour les exercices de mise en route. Les étapes vous indiquent comment configurer votre compte, créer un IAM rôle Amazon Kendra autorisant à passer des appels en votre nom et indexer les documents d'un Amazon S3 bucket. Un compartiment S3 est utilisé à titre d'exemple, mais vous pouvez utiliser d'autres sources de données Amazon Kendra compatibles. Voir Sources de données.

Inscrivez-vous pour un Compte AWS

Si vous n'en avez pas Compte AWS, procédez comme suit pour en créer un.

Pour vous inscrire à un Compte AWS

1. Ouvrez l'https://portal.aws.amazon.com/billing/inscription.
2. Suivez les instructions en ligne.

   Dans le cadre de la procédure d’inscription, vous recevrez un appel téléphonique ou un SMS et vous saisirez un code de vérification en utilisant le clavier numérique du téléphone.

Lorsque vous vous inscrivez à un Compte AWS, un Utilisateur racine d'un compte AWSest créé. Par défaut, seul l'utilisateur racine a accès à l'ensemble des Services AWS et des ressources de ce compte. La meilleure pratique de sécurité consiste à attribuer un accès administratif à un utilisateur, et à utiliser uniquement l'utilisateur racine pour effectuer les tâches nécessitant un accès utilisateur racine.

AWS vous envoie un e-mail de confirmation une fois le processus d'inscription terminé. À tout moment, vous pouvez consulter l'activité actuelle de votre compte et le gérer en accédant à https://aws.amazon.com/et en choisissant Mon compte.

Création d'un utilisateur doté d'un accès administratif

Une fois que vous vous êtes inscrit à un utilisateur administratif Compte AWS, que vous Utilisateur racine d'un compte AWS l'avez sécurisé AWS IAM Identity Center, que vous l'avez activé et que vous en avez créé un, afin de ne pas utiliser l'utilisateur root pour les tâches quotidiennes.

Sécurisez votre Utilisateur racine d'un compte AWS

1. Connectez-vous en AWS Management Consoletant que propriétaire du compte en choisissant Utilisateur root et en saisissant votre adresse Compte AWS e-mail. Sur la page suivante, saisissez votre mot de passe.

   Pour obtenir de l'aide pour vous connecter en utilisant l'utilisateur racine, consultez Connexion en tant qu'utilisateur racine dans le Guide de l'utilisateur Connexion à AWS .
2. Activez l'authentification multifactorielle (MFA) pour votre utilisateur racine.

   Pour obtenir des instructions, voir Activer un périphérique MFA virtuel pour votre utilisateur Compte AWS root (console) dans le guide de l'utilisateur IAM.

Création d'un utilisateur doté d'un accès administratif

1. Activez IAM Identity Center.

   Pour obtenir des instructions, consultez Activation d' AWS IAM Identity Center dans le Guide de l'utilisateur AWS IAM Identity Center .
2. Dans IAM Identity Center, octroyez un accès administratif à un utilisateur.

Pour un didacticiel sur l'utilisation du Répertoire IAM Identity Center comme source d'identité, voir Configurer l'accès utilisateur par défaut Répertoire IAM Identity Center dans le Guide de AWS IAM Identity Center l'utilisateur.

Connexion en tant qu'utilisateur doté d'un accès administratif

- Pour vous connecter avec votre utilisateur IAM Identity Center, utilisez l'URL de connexion qui a été envoyée à votre adresse e-mail lorsque vous avez créé l'utilisateur IAM Identity Center.

  Pour obtenir de l'aide pour vous connecter en utilisant un utilisateur d'IAM Identity Center, consultez la section Connexion au portail AWS d'accès dans le guide de l'Connexion à AWS utilisateur.

Attribution d'un accès à d'autres utilisateurs

1. Dans IAM Identity Center, créez un ensemble d'autorisations qui respecte la bonne pratique consistant à appliquer les autorisations de moindre privilège.

   Pour obtenir des instructions, consultez Création d'un ensemble d'autorisations dans le Guide de l'utilisateur AWS IAM Identity Center .
2. Attribuez des utilisateurs à un groupe, puis attribuez un accès par authentification unique au groupe.

   Pour obtenir des instructions, consultez Ajout de groupes dans le Guide de l'utilisateur AWS IAM Identity Center .

- Si vous utilisez un compartiment S3 contenant des documents à tester Amazon Kendra, créez un compartiment S3 dans la même région que celle que vous utilisez Amazon Kendra. Pour obtenir des instructions, consultez la section Création et configuration d'un compartiment S3 dans le guide de l'utilisateur d'Amazon Simple Storage Service.

  Téléchargez vos documents dans votre compartiment S3. Pour obtenir des instructions, consultez la section Chargement, téléchargement et gestion d'objets dans le guide de l'utilisateur d'Amazon Simple Storage Service.

  Si vous utilisez une autre source de données, vous devez disposer d'un site actif et d'informations d'identification pour vous connecter à la source de données.

Si vous utilisez la console pour commencer, commencez parCommencer à utiliser la Amazon Kendra console.

Amazon Kendra ressources : SDK AWS CLI, console

Certaines autorisations sont requises si vous utilisez la CLI, le SDK ou la console.

Pour l'utiliser Amazon Kendra pour la CLI, le SDK ou la console, vous devez disposer des autorisations nécessaires Amazon Kendra pour créer et gérer des ressources en votre nom. Selon votre cas d'utilisation, ces autorisations incluent l'accès à l' Amazon Kendra API elle-même, AWS KMS keys si vous souhaitez chiffrer vos données via une clé CMK personnalisée, le répertoire Identity Center si vous souhaitez intégrer AWS IAM Identity Center ou créer une expérience de recherche. Pour obtenir la liste complète des autorisations pour différents cas d'utilisation, consultez la section IAM rôles.

Tout d'abord, vous devez associer les autorisations ci-dessous à votre utilisateur IAM.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "Stmt1644430853544",
      "Action": [
        "kms:CreateGrant",
        "kms:DescribeKey"
      ],
      "Effect": "Allow",
      "Resource": "*"
    },
    {
      "Sid": "Stmt1644430878150",
      "Action": "kendra:*",
      "Effect": "Allow",
      "Resource": "*"
    },
    {
      "Sid": "Stmt1644430973706",
      "Action": [
        "sso:AssociateProfile",
        "sso:CreateManagedApplicationInstance",
```

```
        "sso:DeleteManagedApplicationInstance",
        "sso:DisassociateProfile",
        "sso:GetManagedApplicationInstance",
        "sso:GetProfile",
        "sso:ListDirectoryAssociations",
        "sso:ListProfileAssociations",
        "sso:ListProfiles"
      ],
      "Effect": "Allow",
      "Resource": "*"
    },
    {
      "Sid": "Stmt1644430999558",
      "Action": [
        "sso-directory:DescribeGroup",
        "sso-directory:DescribeGroups",
        "sso-directory:DescribeUser",
        "sso-directory:DescribeUsers"
      ],
      "Effect": "Allow",
      "Resource": "*"
    },
    {
      "Sid": "Stmt1644431025960",
      "Action": [
        "identitystore:DescribeGroup",
        "identitystore:DescribeUser",
        "identitystore:ListGroups",
        "identitystore:ListUsers"
      ],
      "Effect": "Allow",
      "Resource": "*"
    }
  ]
}
```

Ensuite, si vous utilisez la CLI ou le SDK, vous devez également créer un IAM rôle et une politique d'accès Amazon CloudWatch Logs. Si vous utilisez la console, il n'est pas nécessaire de créer un IAM rôle ni une politique pour cela. Vous le créez dans le cadre de la procédure de console.

Pour créer un IAM rôle et une politique pour le SDK AWS CLI et permettant d'accéder Amazon Kendra à votre Amazon CloudWatch Logs.

1. Connectez-vous à la console IAM AWS Management Console et ouvrez-la à https://console.aws.amazon.com/iam/l'adresse.
2. Dans le menu de gauche, choisissez Politiques, puis choisissez Créer une politique.
3. Choisissez JSON, puis remplacez la politique par défaut par la suivante :

   JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "cloudwatch:PutMetricData"
            ],
            "Resource": "*",
            "Condition": {
                "StringEquals": {
                    "cloudwatch:namespace": "AWS/Kendra"
                }
            }
        },
        {
            "Effect": "Allow",
            "Action": [
                "logs:DescribeLogGroups"
            ],
            "Resource": "*"
        },
        {
            "Effect": "Allow",
            "Action": [
                "logs:CreateLogGroup"
            ],
            "Resource": [
                "arn:aws:logs:region:account ID:log-group:/aws/kendra/*"
            ]
        },
```

```
        {
            "Effect": "Allow",
            "Action": [
                "logs:DescribeLogStreams",
                "logs:CreateLogStream",
                "logs:PutLogEvents"
            ],
            "Resource": [
                "arn:aws:logs:region:account ID:log-group:/aws/kendra/*:log-
stream:*"
            ]
        }
    ]
}
```

4. Choisissez Examiner une politique.
5. Nommez la politique, "KendraPolicyForGettingStartedIndex" puis choisissez Create policy.
6. Dans le menu de gauche, choisissez Rôles, puis sélectionnez Créer un rôle.
7. Choisissez Un autre AWS compte, puis saisissez votre identifiant de compte dans Identifiant du compte. Choisissez Suivant : Autorisations.
8. Choisissez la politique que vous avez créée ci-dessus, puis cliquez sur Suivant : Tags
9. N'ajoutez aucune balise. Choisissez Suivant : Vérification.
10. Nommez le rôle, "KendraRoleForGettingStartedIndex" puis choisissez Créer un rôle.
11. Trouvez le rôle que vous venez de créer. Choisissez le nom du rôle pour ouvrir le résumé. Choisissez Relations de confiance, puis Modifier la relation de confiance.
12. Remplacez la relation de confiance existante par la suivante :

    JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
      {
        "Effect": "Allow",
        "Principal": {
          "Service": "kendra.amazonaws.com"
        },
        "Action": "sts:AssumeRole"
      }
```

```
    ]
}
```

13. Choisissez Update Trust Policy (Mettre à jour la stratégie d'approbation).

Troisièmement, si vous utilisez un Amazon S3 pour stocker vos documents ou si vous utilisez S3 pour les tester Amazon Kendra, vous devez également créer un IAM rôle et une politique pour accéder à votre compartiment. Si vous utilisez une autre source de données, consultez la section IAM Rôles des sources de données.

Pour créer un IAM rôle et une politique permettant d'accéder Amazon Kendra à votre Amazon S3 compartiment et de l'indexer.

1. Connectez-vous à la console IAM AWS Management Console et ouvrez-la à https://console.aws.amazon.com/iam/l'adresse.
2. Dans le menu de gauche, choisissez Politiques, puis choisissez Créer une politique.
3. Choisissez JSON, puis remplacez la politique par défaut par la suivante :

   JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Action": [
                "s3:GetObject"
            ],
            "Resource": [
                "arn:aws:s3:::bucket name/*"
            ],
            "Effect": "Allow"
        },
        {
            "Action": [
                "s3:ListBucket"
            ],
            "Resource": [
                "arn:aws:s3:::bucket name"
            ],
            "Effect": "Allow"
```

```
        },
        {
            "Effect": "Allow",
            "Action": [
                "kendra:BatchPutDocument",
                "kendra:BatchDeleteDocument"
            ],
            "Resource": "arn:aws:kendra:region:account ID:index/*"
        }
    ]
}
```

4. Choisissez Examiner une politique.
5. Nommez la politique « KendraPolicyForGettingStartedDataSource », puis choisissez Create policy.
6. Dans le menu de gauche, choisissez Rôles, puis sélectionnez Créer un rôle.
7. Choisissez Un autre AWS compte, puis saisissez votre identifiant de compte dans Identifiant du compte. Choisissez Suivant : Autorisations.
8. Choisissez la politique que vous avez créée ci-dessus, puis cliquez sur Suivant : Tags
9. N'ajoutez aucune balise. Choisissez Suivant : Vérification.
10. Nommez le rôle « KendraRoleForGettingStartedDataSource », puis choisissez Créer un rôle.
11. Trouvez le rôle que vous venez de créer. Choisissez le nom du rôle pour ouvrir le résumé. Choisissez Relations de confiance, puis Modifier la relation de confiance.
12. Remplacez la relation de confiance existante par la suivante :

    JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
      {
        "Effect": "Allow",
        "Principal": {
          "Service": "kendra.amazonaws.com"
        },
        "Action": "sts:AssumeRole"
      }
    ]
```

```
}
```

13. Choisissez Update Trust Policy (Mettre à jour la stratégie d'approbation).

Selon la manière dont vous souhaitez utiliser l' Amazon Kendra API, effectuez l'une des opérations suivantes.

- Démarrer (AWS CLI)
- Démarrer (AWS SDK pour Java)
- Démarrer (AWS SDK pour Python (Boto3))

Commencer à utiliser la Amazon Kendra console

Les procédures suivantes montrent comment créer et tester un Amazon Kendra index à l'aide de la AWS console. Dans les procédures, vous créez un index et une source de données pour un index. Enfin, vous testez votre index en effectuant une demande de recherche.

Étape 1 : Pour créer un index (console)

1. Connectez-vous à la console AWS de gestion et ouvrez-la à l' Amazon Kendra adresse https://console.aws.amazon.com/kendra/.
2. Sélectionnez Créer un index dans la section Index.
3. Sur la page Spécifier les détails de l'index, donnez un nom et une description à votre index.
4. Dans IAM rôle, choisissez Créer un nouveau rôle, puis nommez le rôle. Le IAM rôle aura le préfixe « AmazonKendra - ».
5. Conservez les valeurs par défaut de tous les autres champs. Choisissez Suivant.
6. Sur la page Configurer le contrôle d'accès utilisateur, choisissez Next.
7. Sur la page des détails du provisionnement, choisissez Developer Edition.
8. Choisissez Créer pour créer votre index.
9. Attendez que votre index soit créé. Amazon Kendra fournit le matériel nécessaire à votre index. Cette opération peut prendre un certain temps.

Étape 2 : Pour ajouter une source de données à un index (console)

1. Consultez les sources de données disponibles pour vous connecter Amazon Kendra à vos documents et les indexer.
2. Dans le volet de navigation, sélectionnez Sources de données, puis sélectionnez Ajouter une source de données pour la source de données de votre choix.
3. Suivez les étapes pour configurer la source de données.

Étape 3 : Pour rechercher un index (console)

1. Dans le volet de navigation, choisissez l'option permettant d'effectuer une recherche dans votre index.
2. Entrez un terme de recherche adapté à votre index. Les meilleurs résultats et les meilleurs résultats du document sont affichés.

Démarrer (AWS CLI)

La procédure suivante montre comment créer un Amazon Kendra index à l'aide du AWS CLI. La procédure crée une source de données, un index et exécute une requête sur l'index.

Pour créer un Amazon Kendra index (CLI)

1. Fais lePrérequis.
2. Entrez la commande suivante pour créer un index.

```
aws kendra create-index \
 --name cli-getting-started-index \
 --description "Index for CLI getting started guide." \
 --role-arn arn:aws:iam::account id:role/KendraRoleForGettingStartedIndex
```

3. Attendez Amazon Kendra de créer l'index. Vérifiez la progression à l'aide de la commande suivante. Lorsque le champ d'état est `ACTIVE` défini, passez à l'étape suivante.

```
aws kendra describe-index \
 --id index id
```

4. À l'invite de commande, entrez la commande suivante pour créer une source de données.

```
aws kendra create-data-source \
 --index-id index id \
 --name data source name \
 --role-arn arn:aws:iam::account id:role/KendraRoleForGettingStartedDataSource \
 --type S3 \
 --configuration '{"S3Configuration":{"BucketName":"S3 bucket name"}}'
```

Si vous vous connectez à votre source de données à l'aide d'un schéma de modèle, configurez le schéma de modèle.

```
aws kendra create-data-source \
 --index-id index id \
 --name data source name \
 --role-arn arn:aws:iam::account id:role/KendraRoleForGettingStartedDataSource \
 --type TEMPLATE \
 --configuration '{"TemplateConfiguration":{"Template":{JSON schema}}}'
```

5. La création de la source de données prendra Amazon Kendra un certain temps. Entrez la commande suivante pour vérifier la progression. Lorsque le statut est `ACTIVE` défini, passez à l'étape suivante.

```
aws kendra describe-data-source \
 --id data source ID \
 --index-id index ID
```

6. Entrez la commande suivante pour synchroniser la source de données.

```
aws kendra start-data-source-sync-job \
 --id data source ID \
 --index-id index ID
```

7. Amazon Kendra indexera votre source de données. Le temps nécessaire dépend du nombre de documents. Vous pouvez vérifier l'état de la tâche de synchronisation à l'aide de la commande suivante. Lorsque le statut est `ACTIVE` défini, passez à l'étape suivante.

```
aws kendra describe-data-source \
 --id data source ID \
 --index-id index ID
```

8. Entrez la commande suivante pour effectuer une requête.

```
aws kendra query \
 --index-id index ID \
 --query-text "search term"
```

Les résultats de la recherche sont affichés au format JSON.

Démarrer (AWS SDK pour Python (Boto3))

Le programme suivant est un exemple d'utilisation Amazon Kendra dans un programme Python. Le programme exécute les actions suivantes :

1. Crée un nouvel index à l'aide de l'CreateIndexopération.
2. Attend la fin de la création de l'index. Il utilise l'DescribeIndexopération pour surveiller l'état de l'index.
3. Une fois que l'index est actif, il crée une source de données à l'aide de l'CreateDataSourceopération.
4. Attend que la création de la source de données soit terminée. Il utilise l'DescribeDataSourceopération pour surveiller l'état de la source de données.
5. Lorsque la source de données est active, elle synchronise l'index avec le contenu de la source de données à l'aide de l'StartDataSourceSyncJobopération.

```
import boto3
from botocore.exceptions import ClientError
import pprint
import time

kendra = boto3.client("kendra")

print("Create an index.")

# Provide a name for the index
index_name = "python-getting-started-index"
# Provide an optional decription for the index
description = "Getting started index"
# Provide the IAM role ARN required for indexes
index_role_arn = "arn:aws:iam::${accountId}:role/KendraRoleForGettingStartedIndex"
```

```
try:
    index_response = kendra.create_index(
        Description = description,
        Name = index_name,
        RoleArn = index_role_arn
    )

    pprint.pprint(index_response)

    index_id = index_response["Id"]

    print("Wait for Amazon Kendra to create the index.")

    while True:
        # Get the details of the index, such as the status
        index_description = kendra.describe_index(
            Id = index_id
        )
        # When status is not CREATING quit.
        status = index_description["Status"]
        print(" Creating index. Status: "+status)
        time.sleep(60)
        if status != "CREATING":
            break

    print("Create an S3 data source.")

    # Provide a name for the data source
    data_source_name = "python-getting-started-data-source"
    # Provide an optional description for the data source
    data_source_description = "Getting started data source."
    # Provide the IAM role ARN required for data sources
    data_source_role_arn = "arn:aws:iam::${accountId}:role/
KendraRoleForGettingStartedDataSource"
    # Provide the data source connection information
    S3_bucket_name = "S3-bucket-name"
    data_source_type = "S3"
    # Configure the data source
    configuration = {"S3Configuration":
        {
            "BucketName": S3_bucket_name
        }
    }
```

```
    """
    If you connect to your data source using a template schema,
    configure the template schema
    configuration = {"TemplateConfiguration":
        {
            "Template": {JSON schema}
        }
    }
    """

    data_source_response = kendra.create_data_source(
        Name = data_source_name,
        Description = data_source_name,
        RoleArn = data_source_role_arn,
        Type = data_source_type,
        Configuration = configuration,
        IndexId = index_id
    )

    pprint.pprint(data_source_response)

    data_source_id = data_source_response["Id"]

    print("Wait for Amazon Kendra to create the data source.")

    while True:
        # Get the details of the data source, such as the status
        data_source_description = kendra.describe_data_source(
            Id = data_source_id,
            IndexId = index_id
        )
        # If status is not CREATING, then quit
        status = data_source_description["Status"]
        print(" Creating data source. Status: "+status)
        time.sleep(60)
        if status != "CREATING":
            break

    print("Synchronize the data source.")

    sync_response = kendra.start_data_source_sync_job(
        Id = data_source_id,
        IndexId = index_id
    )
```

```
    pprint.pprint(sync_response)

    print("Wait for the data source to sync with the index.")

    while True:

        jobs = kendra.list_data_source_sync_jobs(
            Id = data_source_id,
            IndexId = index_id
        )

        # For this example, there should be one job
        status = jobs["History"][0]["Status"]

        print(" Syncing data source. Status: "+status)
        if status != "SYNCING":
            break
        time.sleep(60)

except  ClientError as e:
        print("%s" % e)

print("Program ends.")
```

Démarrer (AWS SDK pour Java)

Le programme suivant est un exemple d'utilisation Amazon Kendra dans un programme Java. Le programme exécute les actions suivantes :

1. Crée un nouvel index à l'aide de l'CreateIndexopération.
2. Attend la fin de la création de l'index. Il utilise l'DescribeIndexopération pour surveiller l'état de l'index.
3. Une fois que l'index est actif, il crée une source de données à l'aide de l'CreateDataSourceopération.
4. Attend que la création de la source de données soit terminée. Il utilise l'DescribeDataSourceopération pour surveiller l'état de la source de données.
5. Lorsque la source de données est active, elle synchronise l'index avec le contenu de la source de données à l'aide de l'StartDataSourceSyncJobopération.

```
package com.amazonaws.kendra;

import java.util.concurrent.TimeUnit;
import software.amazon.awssdk.services.kendra.KendraClient;
import software.amazon.awssdk.services.kendra.model.CreateDataSourceRequest;
import software.amazon.awssdk.services.kendra.model.CreateDataSourceResponse;
import software.amazon.awssdk.services.kendra.model.CreateIndexRequest;
import software.amazon.awssdk.services.kendra.model.CreateIndexResponse;
import software.amazon.awssdk.services.kendra.model.DataSourceConfiguration;
import software.amazon.awssdk.services.kendra.model.DataSourceStatus;
import software.amazon.awssdk.services.kendra.model.DataSourceSyncJob;
import software.amazon.awssdk.services.kendra.model.DataSourceSyncJobStatus;
import software.amazon.awssdk.services.kendra.model.DataSourceType;
import software.amazon.awssdk.services.kendra.model.DescribeDataSourceRequest;
import software.amazon.awssdk.services.kendra.model.DescribeDataSourceResponse;
import software.amazon.awssdk.services.kendra.model.DescribeIndexRequest;
import software.amazon.awssdk.services.kendra.model.DescribeIndexResponse;
import software.amazon.awssdk.services.kendra.model.IndexStatus;
import software.amazon.awssdk.services.kendra.model.ListDataSourceSyncJobsRequest;
import software.amazon.awssdk.services.kendra.model.ListDataSourceSyncJobsResponse;
import software.amazon.awssdk.services.kendra.model.S3DataSourceConfiguration;
import software.amazon.awssdk.services.kendra.model.StartDataSourceSyncJobRequest;
import software.amazon.awssdk.services.kendra.model.StartDataSourceSyncJobResponse;


public class CreateIndexAndDataSourceExample {

    public static void main(String[] args) throws InterruptedException {
        System.out.println("Create an index");

        String indexDescription = "Getting started index for Kendra";
        String indexName = "java-getting-started-index";
        String indexRoleArn = "arn:aws:iam::<your AWS account ID>:role/<name of an IAM
 role>";

        System.out.println(String.format("Creating an index named %s", indexName));
        KendraClient kendra = KendraClient.builder().build();

        CreateIndexRequest createIndexRequest = CreateIndexRequest
            .builder()
            .description(indexDescription)
            .name(indexName)
            .roleArn(indexRoleArn)
```

```
            .build();
        CreateIndexResponse createIndexResponse =
kendra.createIndex(createIndexRequest);
        System.out.println(String.format("Index response %s", createIndexResponse));

        String indexId = createIndexResponse.id();

        System.out.println(String.format("Waiting until the index with index ID %s is
created", indexId));
        while (true) {
            DescribeIndexRequest describeIndexRequest =
DescribeIndexRequest.builder().id(indexId).build();
            DescribeIndexResponse describeIndexResponse =
kendra.describeIndex(describeIndexRequest);
            IndexStatus status = describeIndexResponse.status();
            if (status != IndexStatus.CREATING) {
                break;
            }

            TimeUnit.SECONDS.sleep(60);
        }

        System.out.println("Creating an S3 data source");
        String dataSourceName = "java-getting-started-data-source";
        String dataSourceDescription = "Getting started data source";
        String s3BucketName = "amzn-s3-demo-bucket";
        String dataSourceRoleArn = "arn:aws:iam::<your AWS account ID>:role/<name of an
IAM role>";

        CreateDataSourceRequest createDataSourceRequest = CreateDataSourceRequest
            .builder()
            .indexId(indexId)
            .name(dataSourceName)
            .description(dataSourceDescription)
            .roleArn(dataSourceRoleArn)
            .type(DataSourceType.S3)
            .configuration(
                DataSourceConfiguration
                    .builder()
                    .s3Configuration(
                        S3DataSourceConfiguration
                            .builder()
                            .bucketName(s3BucketName)
                            .build()
```

```
                    ).build()
            ).build();

        CreateDataSourceResponse createDataSourceResponse =
kendra.createDataSource(createDataSourceRequest);
        System.out.println(String.format("Response of creating data source: %s",
createDataSourceResponse));

        String dataSourceId = createDataSourceResponse.id();
        System.out.println(String.format("Waiting for Kendra to create the data source
%s", dataSourceId));
        DescribeDataSourceRequest describeDataSourceRequest = DescribeDataSourceRequest
            .builder()
            .indexId(indexId)
            .id(dataSourceId)
            .build();

        while (true) {
            DescribeDataSourceResponse describeDataSourceResponse =
kendra.describeDataSource(describeDataSourceRequest);

            DataSourceStatus status = describeDataSourceResponse.status();
            System.out.println(String.format("Creating data source. Status: %s",
status));
            if (status != DataSourceStatus.CREATING) {
                break;
            }

            TimeUnit.SECONDS.sleep(60);
        }

        System.out.println(String.format("Synchronize the data source %s",
dataSourceId));
        StartDataSourceSyncJobRequest startDataSourceSyncJobRequest =
StartDataSourceSyncJobRequest
            .builder()
            .indexId(indexId)
            .id(dataSourceId)
            .build();
        StartDataSourceSyncJobResponse startDataSourceSyncJobResponse =
kendra.startDataSourceSyncJob(startDataSourceSyncJobRequest);
        System.out.println(String.format("Waiting for the data
source to sync with the index %s for execution ID %s", indexId,
startDataSourceSyncJobResponse.executionId()));
```

```
        // For this particular list, there should be just one job
        ListDataSourceSyncJobsRequest listDataSourceSyncJobsRequest =
 ListDataSourceSyncJobsRequest
            .builder()
            .indexId(indexId)
            .id(dataSourceId)
            .build();

        while (true) {
            ListDataSourceSyncJobsResponse listDataSourceSyncJobsResponse =
 kendra.listDataSourceSyncJobs(listDataSourceSyncJobsRequest);
            DataSourceSyncJob job = listDataSourceSyncJobsResponse.history().get(0);
            System.out.println(String.format("Syncing data source. Status: %s",
 job.status()));

            if (job.status() != DataSourceSyncJobStatus.SYNCING) {
                break;
            }

            TimeUnit.SECONDS.sleep(60);

        }

        System.out.println("Index setup is complete");
    }
}
```

Commencer à utiliser une source de Amazon S3 données (console)

Vous pouvez utiliser la Amazon Kendra console pour commencer à utiliser un Amazon S3 bucket comme magasin de données. Lorsque vous utilisez la console, vous spécifiez toutes les informations de connexion dont vous avez besoin pour indexer le contenu du bucket. Pour de plus amples informations, veuillez consulter Amazon S3.

Utilisez la procédure suivante pour créer une source de données de compartiment S3 de base à l'aide de la configuration par défaut. La procédure suppose que vous avez créé un index en suivant les étapes de l'étape 1 deCommencer à utiliser la Amazon Kendra console.

Pour créer une source de données de compartiment S3 à l'aide de la Amazon Kendra console

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la à la https://console.aws.amazon.com/kendra/maison.
2. Dans la liste des index, choisissez l'index auquel vous souhaitez ajouter la source de données.
3. Choisissez Ajouter des sources de données.
4. Dans la liste des connecteurs de source de données, sélectionnez Amazon S3.
5. Sur la page Définir les attributs, donnez un nom à votre source de données et éventuellement une description. Laissez le champ Tags vide. Choisissez Next (Suivant) pour continuer.
6. Dans le champ Entrez l'emplacement de la source de données, entrez le nom du compartiment S3 qui contient vos documents. Vous pouvez saisir le nom directement ou rechercher le nom en choisissant Parcourir. Le compartiment doit se trouver dans la même région que l'index.
7. Dans IAM Rôle, choisissez Créer un nouveau rôle, puis tapez un nom de rôle. Pour plus d'informations, consultez la section IAM Rôles des sources de Amazon S3 données.
8. Dans la section Définir le calendrier de synchronisation, choisissez Exécuter à la demande.
9. Choisissez Next (Suivant) pour continuer.
10. Sur la page Réviser et créer, passez en revue les détails de votre source de données S3. Si vous souhaitez apporter des modifications, cliquez sur le bouton Modifier situé à côté de l'élément que vous souhaitez modifier. Lorsque vous êtes satisfait de vos choix, choisissez Create pour créer votre source de données S3.

Après avoir sélectionné Create, Amazon Kendra commence à créer la source de données. La création de la source de données peut prendre plusieurs minutes. Une fois l'opération terminée, le statut de la source de données passe de Création à Actif.

Après avoir créé la source de données, vous devez synchroniser l' Amazon Kendra index avec la source de données. Choisissez Synchroniser maintenant pour démarrer le processus de synchronisation. La synchronisation de la source de données peut prendre de quelques minutes à plusieurs heures, selon le nombre et la taille des documents.

Commencer à utiliser une source de données de base de données MySQL (console)

Vous pouvez utiliser la Amazon Kendra console pour commencer à utiliser une base de données MySQL comme source de données. Lorsque vous utilisez la console, vous spécifiez les informations

de connexion dont vous avez besoin pour indexer le contenu d'une base de données MySQL. Pour plus d'informations, consultez Utilisation d'une source de données de base de données.

Vous devez d'abord créer une base de données MySQL, puis vous pouvez créer une source de données pour la base de données.

Utilisez la procédure suivante pour créer une base de données MySQL de base. La procédure suppose que vous avez déjà créé un index après l'étape 1 deCommencer à utiliser la Amazon Kendra console.

Pour créer une base de données MySQL

1. Connectez-vous à la console Amazon RDS AWS Management Console et ouvrez-la à https://console.aws.amazon.com/rds/l'adresse.
2. Dans le volet de navigation, choisissez Subnet groups, puis Create DB Subnet Group.
3. Nommez le groupe et choisissez votre Virtual Private Cloud (VPC). Pour plus d'informations sur la configuration d'un VPC, consultez Configuration Amazon Kendra pour utiliser un VPC.
4. Ajoutez les sous-réseaux privés de votre VPC. Vos sous-réseaux privés sont ceux qui ne sont pas connectés à votre NAT. Choisissez Créer.
5. Dans le volet de navigation, choisissez Databases, puis Create database.
6. Utilisez les paramètres suivants pour créer la base de données. Conservez les valeurs par défaut de tous les autres paramètres.
   - Options du moteur —MySQL
   - Modèles — Niveau gratuit
   - Paramètres d'identification —Entrez et confirmez un mot de passe
   - Sous Connectivité, sélectionnez Configuration de connectivité supplémentaire. Effectuez les choix suivants.
     - Groupe de sous-réseaux : choisissez le groupe de sous-réseaux que vous avez créé à l'étape 4.
     - Groupe de sécurité VPC : choisissez le groupe contenant les règles entrantes et sortantes que vous avez créées dans votre VPC. Par exemple, **`DataSourceSecurityGroup`**. Pour plus d'informations sur la configuration d'un VPC, consultez Configuration Amazon Kendra pour utiliser un VPC.
   - Sous Configuration supplémentaire, définissez le nom de base de données initial sur**`content`**.
7. Choisissez Créer une base de données.

8. Dans la liste des bases de données, choisissez votre nouvelle base de données. Notez le point de terminaison de la base de données.
9. Après avoir créé votre base de données, vous devez créer une table contenant vos documents. La création d'une table n'entre pas dans le cadre de ces instructions. Lorsque vous créez votre tableau, tenez compte des points suivants :

    - Nom de la base de données : **`content`**
    - Nom de la table : **`documents`**
    - Colonnes— **`IDTitle`**,**`Body`**, et**`LastUpdate`**. Vous pouvez inclure des colonnes supplémentaires si vous le souhaitez.

Maintenant que vous avez créé votre base de données MySQL, vous pouvez créer une source de données pour la base de données.

Pour créer une source de données MySQL

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la à la https://console.aws.amazon.com/kendra/maison.
2. Dans le volet de navigation, choisissez Indexes, puis choisissez votre index.
3. Choisissez Ajouter des sources de données, puis Amazon RDS.
4. Tapez le nom et la description de la source de données, puis choisissez Next.
5. Choisissez MySQL.
6. Sous Accès à la connexion, entrez les informations suivantes :

    - Point de terminaison : point de terminaison de la base de données que vous avez créée précédemment.
    - Port : numéro de port de la base de données. Pour MySQL, la valeur par défaut est 3306.
    - Type d'authentification —Choisissez Nouveau.
    - Nouveau nom de conteneur secret : nom du Secrets Manager conteneur pour les informations d'identification de base de données.
    - Nom d'utilisateur : nom d'un utilisateur disposant d'un accès administratif à la base de données.
    - Mot de passe : mot de passe de l'utilisateur, puis choisissez Enregistrer l'authentification.
    - Nom de la base de données —**`content`**.

- Nom de la table —**`documents`**.
- Rôle IAM —Choisissez Créer un nouveau rôle, puis tapez le nom du rôle.

7. Dans Configuration des colonnes, entrez les informations suivantes :
   - Nom de la colonne d'identification du document — **`ID`**
   - Nom de la colonne du titre du document — **`Title`**
   - Nom de la colonne de données du document : **`Body`**
8. Dans Détection des changements de colonne, entrez les informations suivantes :
   - Colonnes de détection des modifications — **`LastUpdate`**
9. Dans Configurer le VPC et le groupe de sécurité, fournissez les informations suivantes :
   - Dans Virtual Private Cloud (VPC), choisissez votre VPC.
   - Dans Sous-réseaux, choisissez les sous-réseaux privés que vous avez créés dans votre VPC.
   - Dans les groupes de sécurité VPC, choisissez le groupe de sécurité qui contient les règles entrantes et sortantes que vous avez créées dans votre VPC pour les bases de données MySQL. Par exemple, **`DataSourceSecurityGroup`**.
10. Dans Définir le calendrier de synchronisation, choisissez Exécuter à la demande, puis Suivant.
11. Dans Mappage des champs de source de données, choisissez Next.
12. Vérifiez la configuration de votre source de données pour vous assurer qu'elle est correcte. Lorsque vous êtes certain que tout est correct, choisissez Create.

Commencer à utiliser une source AWS IAM Identity Center d'identité (console)

Une source AWS IAM Identity Center d'identité contient des informations sur vos utilisateurs et vos groupes. Cela est utile pour configurer le filtrage du contexte utilisateur, qui Amazon Kendra filtre les résultats de recherche pour différents utilisateurs en fonction de l'accès de l'utilisateur ou de son groupe aux documents.

Pour créer une source d'identité IAM Identity Center, vous devez activer IAM Identity Center et créer une organisation dans. AWS Organizations Lorsque vous activez IAM Identity Center et que vous créez une organisation pour la première fois, le répertoire Identity Center est automatiquement sélectionné par défaut comme source d'identité. Vous pouvez passer à Active Directory (géré par Amazon ou autogéré) ou à un fournisseur d'identité externe comme source d'identité. Pour cela,

vous devez suivre les instructions appropriées : voir Modification de la source d'identité de votre IAM Identity Center. Vous ne pouvez avoir qu'une seule source d'identité par organisation.

Pour que différents niveaux d'accès aux documents soient attribués à vos utilisateurs et à vos groupes, vous devez les inclure dans votre liste de contrôle d'accès lorsque vous ingérez des documents dans votre index. Cela permet à vos utilisateurs et à vos groupes de rechercher des documents Amazon Kendra en fonction de leur niveau d'accès. Lorsque vous émettez une requête, l'ID utilisateur doit correspondre exactement au nom d'utilisateur dans IAM Identity Center.

Vous devez également accorder les autorisations requises pour utiliser IAM Identity Center avec Amazon Kendra. Pour plus d'informations, consultez la section IAM Rôles d'IAM Identity Center.

Pour configurer une source d'identité IAM Identity Center

1. Ouvrez la console IAM Identity Center.
2. Choisissez Enable IAM Identity Center, puis sélectionnez Create AWS organization.

   Le répertoire Identity Center est créé par défaut et un e-mail vous est envoyé pour vérifier l'adresse e-mail associée à l'organisation.
3. Pour ajouter un groupe à votre AWS organisation, dans le volet de navigation, sélectionnez Groups.
4. Sur la page Groupes, choisissez Créer un groupe et entrez le nom et la description du groupe dans la boîte de dialogue. Choisissez Créer.
5. Pour ajouter un utilisateur à vos Organizations, dans le volet de navigation, sélectionnez Users.
6. Sur la page Users (Utilisateurs), choisissez Add user (Ajouter un utilisateur). Sous User details (Détails de l'utilisateur), renseignez tous les champs obligatoires. Pour Password (Mot de passe), choisissez Send an email to the user (Envoyer un e-mail à l'utilisateur). Choisissez Suivant.
7. Pour ajouter un utilisateur à un groupe, choisissez Groupes, puis sélectionnez un groupe.
8. Sur la page Détails, sous Membres du groupe, choisissez Ajouter un utilisateur.
9. Sur la page Ajouter des utilisateurs au groupe, sélectionnez l'utilisateur que vous souhaitez ajouter en tant que membre du groupe. Vous pouvez sélectionner plusieurs utilisateurs à ajouter à un groupe.
10. Pour synchroniser votre liste d'utilisateurs et de groupes avec IAM Identity Center, remplacez votre source d'identité par Active Directory ou fournisseur d'identité externe.

    Le répertoire Identity Center est la source d'identité par défaut et vous oblige à ajouter manuellement vos utilisateurs et groupes à l'aide de cette source si vous ne disposez pas de

votre propre liste gérée par un fournisseur. Pour modifier votre source d'identité, vous devez suivre les instructions appropriées : voir Modification de votre source d'identité IAM Identity Center.

> **Note**
>
> Si vous utilisez Active Directory ou un fournisseur d'identité externe comme source d'identité, vous devez associer les adresses e-mail de vos utilisateurs aux noms d'utilisateur IAM Identity Center lorsque vous spécifiez le protocole SCIM (System for Cross-domain Identity Management). Pour plus d'informations, consultez le guide IAM Identity Center sur SCIM pour activer IAM Identity Center.

Une fois que vous avez configuré votre source d'identité IAM Identity Center, vous pouvez l'activer dans la console lorsque vous créez ou modifiez votre index. Accédez à Contrôle d'accès utilisateur dans les paramètres de votre index et modifiez vos paramètres pour autoriser la récupération des informations sur les groupes d'utilisateurs depuis IAM Identity Center.

Vous pouvez également activer IAM Identity Center à l'aide de l'UserGroupResolutionConfigurationobjet. Vous fournissez le `UserGroupResolutionMode` as `AWS_SSO` et créez un IAM rôle qui autorise à appeler `sso:ListDirectoryAssociationssso-directory:SearchUsers`,`sso-directory:ListGroupsForUser`,`sso-directory:DescribeGroups`.

> **Warning**
>
> Amazon Kendra ne prend actuellement pas en charge l'utilisation `UserGroupResolutionConfiguration` avec un compte membre de AWS l'organisation pour votre source d'identité IAM Identity Center. Vous devez créer votre index dans le compte de gestion de l'organisation pour pouvoir l'utiliser`UserGroupResolutionConfiguration`.

Vous trouverez ci-dessous un aperçu de la configuration d'une source de données avec un contrôle `UserGroupResolutionConfiguration` d'accès utilisateur pour filtrer les résultats de recherche en fonction du contexte utilisateur. Cela suppose que vous avez déjà créé un index et un IAM rôle pour les index. Vous créez un index et fournissez le IAM rôle à l'aide de l'CreateIndexAPI.

Configuration d'une source de données avec **`UserGroupResolutionConfiguration`** filtrage du contexte utilisateur

1. Créez un IAM rôle qui autorise l'accès à votre source d'identité IAM Identity Center.
2. Configurez `UserGroupResolutionConfiguration`en définissant le mode `AWS_SSO` et en appelant UpdateIndexpour mettre à jour votre index afin d'utiliser IAM Identity Center.
3. Si vous souhaitez utiliser le contrôle d'accès utilisateur basé sur des jetons pour filtrer les résultats de recherche en fonction du contexte utilisateur, définissez cette option UserContextPolicy`USER_TOKEN`lorsque vous appelez. `UpdateIndex` Sinon, Amazon Kendra parcourt la liste de contrôle d'accès de chacun de vos documents pour la plupart des connecteurs de source de données. Vous pouvez également filtrer les résultats de recherche en fonction du contexte utilisateur dans l'API Query en fournissant des informations sur les utilisateurs et les groupes dans`UserContext`. Vous pouvez également associer les utilisateurs à leurs groupes à l'aide de PutPrincipalMappingtelle sorte que vous n'ayez à fournir l'ID utilisateur que lorsque vous émettez la requête.
4. Créez un IAM rôle qui autorise l'accès à votre source de données.
5. Configurez votre source de données. Vous devez fournir les informations de connexion requises pour vous connecter à votre source de données.
6. Créez une source de données à l'aide de l'CreateDataSourceAPI. Indiquez l'`DataSourceConfiguration`objet, qui inclut `TemplateConfiguration` l'ID de votre index, le IAM rôle de votre source de données, le type de source de données, et nommez votre source de données. Vous pouvez également mettre à jour votre source de données.

Modification de la source d'identité de votre IAM Identity Center

> **Warning**
>
> La modification de votre source d'identité dans les paramètres du centre d'identité IAM peut affecter la conservation des informations relatives aux utilisateurs et aux groupes. Pour le faire en toute sécurité, il est recommandé de consulter les considérations relatives à la modification de votre source d'identité. Lorsque vous modifiez votre source d'identité, un nouvel ID de source d'identité est généré. Vérifiez que vous utilisez le bon identifiant avant de régler le mode sur `AWS_SSO` Activé UserGroupResolutionConfiguration.

Pour modifier la source d'identité de votre IAM Identity Center

1. Ouvrez la console IAM Identity Center>.
2. Sélectionnez Paramètres.
3. Sur la page Paramètres, sous Source d'identité, choisissez Modifier.
4. Sur la page Modifier la source d'identité, sélectionnez votre source d'identité préférée, puis cliquez sur Suivant.

Création d'un index

Vous pouvez créer un index à l'aide de la console ou en appelant l'CreateIndexAPI. Vous pouvez utiliser le AWS Command Line Interface (AWS CLI) ou le SDK avec l'API. Après avoir créé votre index, vous pouvez y ajouter des documents directement ou à partir d'une source de données.

Pour créer un index, vous devez fournir le nom de ressource Amazon (ARN) d'un rôle AWS Identity and Access Management (IAM) auquel les index peuvent accéder CloudWatch. Pour plus d'informations, consultez la section IAM Rôles pour les index.

Les onglets suivants fournissent une procédure pour créer un index à l'aide de AWS Management Console, ainsi que des exemples de code pour utiliser le AWS CLI, et Python et Java SDKs.

Console

Pour créer un index

1. Connectez-vous à la console AWS de gestion et ouvrez-la à l' Amazon Kendra adresse https://console.aws.amazon.com/kendra/.
2. Sélectionnez Créer un index dans la section Index.
3. Dans Spécifier les détails de l'index, donnez un nom et une description à votre index.
4. Dans IAM le rôle, indiquez un IAM rôle. Pour trouver un rôle, choisissez parmi les rôles de votre compte qui contiennent le mot « kendra » ou entrez le nom d'un autre rôle. Pour plus d'informations sur les autorisations requises par le rôle, consultez la section IAM Rôles pour les index.
5. Choisissez Suivant.
6. Sur la page Configurer le contrôle d'accès utilisateur, choisissez Next. Vous pouvez mettre à jour votre index afin d'utiliser des jetons pour le contrôle d'accès après avoir créé un index. Pour plus d'informations, consultez la section Contrôle de l'accès aux documents.
7. Sur la page des détails du provisionnement, choisissez Create.
8. La création de l'index peut prendre un certain temps. Consultez la liste des index pour suivre la progression de la création de votre index. Lorsque le statut de l'index est défini`ACTIVE`, votre index est prêt à être utilisé.

AWS CLI

Pour créer un index

1. Utilisez la commande suivante pour créer un index. `role-arn`Il doit s'agir du Amazon Resource Name (ARN) d'un IAM rôle capable d'exécuter Amazon Kendra des actions. Pour plus d'informations, consultez la section IAM Rôles.

   La commande est formatée pour Linux et macOS. Si vous utilisez Windows, remplacez le caractère de continuation de ligne Unix (\) par un curseur (^).

   ```
   aws kendra create-index \
    --name index name \
    --description "index description" \
    --role-arn arn:aws:iam::account ID:role/role name
   ```

2. La création de l'index peut prendre un certain temps. Pour vérifier l'état de votre index, utilisez l'ID d'index renvoyé `create-index` par la commande suivante. Lorsque le statut de l'index est défini`ACTIVE`, votre index est prêt à être utilisé.

   ```
   aws kendra describe-index \
    --index-id index ID
   ```

Python

Pour créer un index

- Fournissez des valeurs pour les variables suivantes dans l'exemple de code ci-dessous :
  - `description`: une description de l'index que vous êtes en train de créer. Ce nom est facultatif.
  - `index_name`: le nom de l'index que vous créez.
  - `role_arn`—Le nom de ressource Amazon (ARN) d'un rôle qui peut être exécuté Amazon Kendra APIs. Pour plus d'informations, consultez la section IAM Rôles.

  ```
  import boto3
  from botocore.exceptions import ClientError
  import pprint
  ```

```
import time

kendra = boto3.client("kendra")

print("Create an index.")

# Provide a name for the index
index_name = "index-name"
# Provide an optional description for the index
description = "index description"
# Provide the IAM role ARN required for indexes
role_arn = "arn:aws:iam::${account id}:role/${role name}"

try:
    index_response = kendra.create_index(
        Name = index_name,
        Description = description,
        RoleArn = role_arn
    )

    pprint.pprint(index_response)

    index_id = index_response["Id"]

    print("Wait for Amazon Kendra to create the index.")

    while True:
        # Get the details of the index, such as the status
        index_description = kendra.describe_index(
            Id = index_id
        )
        # If status is not CREATING, then quit
        status = index_description["Status"]
        print(" Creating index. Status: "+status)
        if status != "CREATING":
            break
        time.sleep(60)

except  ClientError as e:
        print("%s" % e)

print("Program ends.")
```

Java

Pour créer un index

- Fournissez des valeurs pour les variables suivantes dans l'exemple de code ci-dessous :
  - `description`: une description de l'index que vous êtes en train de créer. Ce nom est facultatif.
  - `index_name`: le nom de l'index que vous créez.
  - `role_arn`—Le nom de ressource Amazon (ARN) d'un rôle qui peut être exécuté Amazon Kendra APIs. Pour plus d'informations, consultez la section IAM Rôles.

```
package com.amazonaws.kendra;

import java.util.concurrent.TimeUnit;
import software.amazon.awssdk.services.kendra.KendraClient;
import software.amazon.awssdk.services.kendra.model.CreateIndexRequest;
import software.amazon.awssdk.services.kendra.model.CreateIndexResponse;
import software.amazon.awssdk.services.kendra.model.DescribeIndexRequest;
import software.amazon.awssdk.services.kendra.model.DescribeIndexResponse;
import software.amazon.awssdk.services.kendra.model.IndexStatus;


public class CreateIndexExample {

    public static void main(String[] args) throws InterruptedException {

        String indexDescription = "Getting started index for Kendra";
        String indexName = "java-getting-started-index";
        String indexRoleArn = "arn:aws:iam::<your AWS account ID>:role/
KendraRoleForGettingStartedIndex";

        System.out.println(String.format("Creating an index named %s",
 indexName));
        CreateIndexRequest createIndexRequest = CreateIndexRequest
            .builder()
            .description(indexDescription)
            .name(indexName)
            .roleArn(indexRoleArn)
            .build();
        KendraClient kendra = KendraClient.builder().build();
```

```
        CreateIndexResponse createIndexResponse =
 kendra.createIndex(createIndexRequest);
        System.out.println(String.format("Index response %s",
 createIndexResponse));

        String indexId = createIndexResponse.id();

        System.out.println(String.format("Waiting until the index with ID %s is
 created.", indexId));
        while (true) {
            DescribeIndexRequest describeIndexRequest =
 DescribeIndexRequest.builder().id(indexId).build();
            DescribeIndexResponse describeIndexResponse =
 kendra.describeIndex(describeIndexRequest);
            IndexStatus status = describeIndexResponse.status();
            if (status != IndexStatus.CREATING) {
                break;
            }

            TimeUnit.SECONDS.sleep(60);
        }

        System.out.println("Index creation is complete.");
    }
}
```

Après avoir créé votre index, vous y ajoutez des documents. Vous pouvez les ajouter directement ou créer une source de données qui met régulièrement à jour votre index.

Rubriques

- Ajouter des documents directement à un index avec téléchargement par lots
- Ajouter des questions fréquemment posées (FAQs) à un index
- Création de champs de document personnalisés
- Contrôle de l'accès des utilisateurs aux documents à l'aide de jetons

Ajouter des documents directement à un index avec téléchargement par lots

Vous pouvez ajouter des documents directement à un index à l'aide de l'BatchPutDocumentAPI. Vous ne pouvez pas ajouter de documents directement à l'aide de la console. Si vous utilisez la console, vous vous connectez à une source de données pour ajouter des documents à votre index. Les documents peuvent être ajoutés à partir d'un compartiment S3 ou fournis sous forme de données binaires. Pour une liste des types de documents pris en charge par la Amazon Kendra section Types de documents.

L'ajout de documents à un index en utilisant `BatchPutDocument` est une opération asynchrone. Après avoir appelé l'`BatchPutDocument`API, vous l'BatchGetDocumentStatusutilisez pour suivre la progression de l'indexation de vos documents. Lorsque vous appelez l'`BatchGetDocumentStatus`API avec une liste de documents IDs, elle renvoie le statut du document. Lorsque le statut du document est `INDEXED` ou`FAILED`, le traitement du document est terminé. Lorsque le statut est défini`FAILED`, l'`BatchGetDocumentStatus`API renvoie la raison pour laquelle le document n'a pas pu être indexé.

Si vous souhaitez modifier le contenu et les champs ou attributs des métadonnées du document pendant le processus d'ingestion du document, consultez la section Enrichissement Amazon Kendra personnalisé des documents. Si vous souhaitez utiliser une source de données personnalisée, chaque document que vous soumettez à l'aide de l'`BatchPutDocument`API nécessite un ID de source de données et un ID d'exécution sous forme d'attributs ou de champs. Pour plus d'informations, consultez la section Attributs obligatoires pour les sources de données personnalisées.

> **Note**
>
> Chaque identifiant de document doit être unique par index. Vous ne pouvez pas créer de source de données pour indexer vos documents avec leur caractère unique, IDs puis utiliser l'`BatchPutDocument`API pour indexer les mêmes documents, ou vice versa. Vous pouvez supprimer une source de données, puis utiliser l'`BatchPutDocument`API pour indexer les mêmes documents, ou vice versa. L'utilisation du connecteur `BatchPutDocument` et `BatchDeleteDocument` APIs en combinaison avec un connecteur de source de Amazon Kendra données pour le même ensemble de documents peut entraîner des incohérences dans vos données. Nous vous recommandons plutôt d'utiliser le connecteur de source de données Amazon Kendra personnalisé.

Les documents suivants du guide du développeur montrent comment ajouter des documents directement à un index.

Rubriques

- Ajouter des documents à l'aide de l' BatchPutDocumentAPI
- Ajouter des documents à partir d'un compartiment S3

Ajouter des documents à l'aide de l' BatchPutDocumentAPI

L'exemple suivant ajoute un blob de texte à un index en appelant BatchPutDocument. Vous pouvez utiliser l'`BatchPutDocument`API pour ajouter des documents directement à votre index. Pour une liste des types de documents pris en charge par la Amazon Kendra section Types de documents.

Pour un exemple de création d'un index à l'aide du AWS CLI et SDKs, consultez la section Création d'un index. Pour configurer la CLI SDKs, reportez-vous à la section Configuration Amazon Kendra.

> **Note**
>
> Les fichiers ajoutés à l'index doivent se trouver dans un flux d'octets codé en UTF-8.

Dans les exemples suivants, du texte codé en UTF-8 est ajouté à l'index.

CLI

Dans le AWS Command Line Interface, utilisez la commande suivante. La commande est formatée pour Linux et macOS. Si vous utilisez Windows, remplacez le caractère de continuation de ligne Unix (\) par un curseur (^).

```
aws kendra batch-put-document \
   --index-id index-id \
   --documents '{"Id":"doc-id-1", "Blob":"Amazon.com is an online retailer.",
 "ContentType":"PLAIN_TEXT", "Title":"Information about Amazon.com"}'
```

Python

```
import boto3

kendra = boto3.client("kendra")
```

```
# Provide the index ID
index_id = "index-id"

# Provide the title and text
title = "Information about Amazon.com"
text = "Amazon.com is an online retailer."

document = {
    "Id": "1",
    "Blob": text,
    "ContentType": "PLAIN_TEXT",
    "Title": title
}

documents = [
    document
]

result = kendra.batch_put_document(
    IndexId = index_id,
    Documents = documents
)

print(result)
```

Java

```
package com.amazonaws.kendra;


import software.amazon.awssdk.core.SdkBytes;
import software.amazon.awssdk.services.kendra.KendraClient;
import software.amazon.awssdk.services.kendra.model.BatchPutDocumentRequest;
import software.amazon.awssdk.services.kendra.model.BatchPutDocumentResponse;
import software.amazon.awssdk.services.kendra.model.ContentType;
import software.amazon.awssdk.services.kendra.model.Document;

public class AddDocumentsViaAPIExample {
    public static void main(String[] args) {
        KendraClient kendra = KendraClient.builder().build();

        String indexId = "yourIndexId";
```

```
        Document testDoc = Document
            .builder()
            .title("The title of your document")
            .id("a_doc_id")
            .blob(SdkBytes.fromUtf8String("your text content"))
            .contentType(ContentType.PLAIN_TEXT)
            .build();

        BatchPutDocumentRequest batchPutDocumentRequest = BatchPutDocumentRequest
            .builder()
            .indexId(indexId)
            .documents(testDoc)
            .build();

        BatchPutDocumentResponse result =
 kendra.batchPutDocument(batchPutDocumentRequest);

        System.out.println(String.format("BatchPutDocument Result: %s", result));
    }
}
```

Ajouter des documents à partir d'un compartiment S3

Vous pouvez ajouter des documents directement à votre index à partir d'un Amazon S3 bucket à l'aide de l'BatchPutDocumentAPI. Vous pouvez ajouter jusqu'à 10 documents au cours d'un même appel. Lorsque vous utilisez un compartiment S3, vous devez fournir un IAM rôle autorisé à accéder au compartiment contenant vos documents. Vous spécifiez le rôle dans le `RoleArn` paramètre.

L'utilisation de l'BatchPutDocumentAPI pour ajouter des documents à partir d'un Amazon S3 bucket ne s'effectue qu'une seule fois. Pour synchroniser un index avec le contenu d'un bucket, créez une source de Amazon S3 données. Pour plus d'informations, consultez la section Source de Amazon S3 données.

Pour un exemple de création d'un index à l'aide du AWS CLI et SDKs, consultez la section Création d'un index. Pour configurer la CLI SDKs, reportez-vous à la section Configuration Amazon Kendra. Pour plus d'informations sur la création d'un compartiment S3, consultez Amazon Simple Storage Service la documentation.

Dans l'exemple suivant, deux documents Microsoft Word sont ajoutés à l'index à l'aide de l'`BatchPutDocument`API.

Python

```
import boto3

kendra = boto3.client("kendra")

# Provide the index ID
index_id = "index-id"
# Provide the IAM role ARN required to index documents in an S3 bucket
role_arn = "arn:aws:iam::${acccountID}:policy/${roleName}"

doc1_s3_file_data = {
    "Bucket": "bucket-name",
    "Key": "document1.docx"
}

doc1_document = {
    "S3Path": doc1_s3_file_data,
    "Title": "Document 1 title",
    "Id": "doc_1"
}

doc2_s3_file_data = {
    "Bucket": "bucket-name",
    "Key": "document2.docx"
}

doc2_document = {
    "S3Path": doc2_s3_file_data,
    "Title": "Document 2 title",
    "Id": "doc_2"
}

documents = [
    doc1_document,
    doc2_document
]

result = kendra.batch_put_document(
    Documents = documents,
    IndexId = index_id,
    RoleArn = role_arn
)
```

```
print(result)
```

Java

```
package com.amazonaws.kendra;

import software.amazon.awssdk.services.kendra.KendraClient;
import software.amazon.awssdk.services.kendra.model.BatchPutDocumentRequest;
import software.amazon.awssdk.services.kendra.model.BatchPutDocumentResponse;
import software.amazon.awssdk.services.kendra.model.Document;
import software.amazon.awssdk.services.kendra.model.S3Path;

public class AddFilesFromS3Example {
    public static void main(String[] args) {
        KendraClient kendra = KendraClient.builder().build();

        String indexId = "yourIndexId";
        String roleArn = "yourIndexRoleArn";

        Document pollyDoc = Document
            .builder()
            .s3Path(
                S3Path.builder()
                .bucket("amzn-s3-demo-bucket")
                .key("What is Amazon Polly.docx")
                .build())
            .title("What is Amazon Polly")
            .id("polly_doc_1")
            .build();

        Document rekognitionDoc = Document
            .builder()
            .s3Path(
                S3Path.builder()
                .bucket("amzn-s3-demo-bucket")
                .key("What is Amazon Rekognition.docx")
                .build())
            .title("What is Amazon rekognition")
            .id("rekognition_doc_1")
            .build();

        BatchPutDocumentRequest batchPutDocumentRequest = BatchPutDocumentRequest
            .builder()
```

```
            .indexId(indexId)
            .roleArn(roleArn)
            .documents(pollyDoc, rekognitionDoc)
            .build();

        BatchPutDocumentResponse result =
 kendra.batchPutDocument(batchPutDocumentRequest);

        System.out.println(String.format("BatchPutDocument result: %s", result));
    }
}
```

Ajouter des questions fréquemment posées (FAQs) à un index

> **Note**
>
> La prise en charge des fonctionnalités varie en fonction du type d'index et de l'API de recherche utilisés. Pour savoir si cette fonctionnalité est prise en charge pour le type d'index et l'API de recherche que vous utilisez, consultez la section Types d'index.

Vous pouvez ajouter des questions fréquemment posées (FAQs) directement à votre index à l'aide de la console ou de l'CreateFaqAPI. L'ajout FAQs à un index est une opération asynchrone. Vous placez les données de la FAQ dans un fichier que vous stockez dans un Amazon Simple Storage Service bucket. Vous pouvez utiliser des fichiers CSV ou JSON comme entrée pour votre FAQ :

- CSV de base : fichier CSV dans lequel chaque ligne contient une question, une réponse et un URI source facultatif.
- CSV personnalisé : fichier CSV contenant des questions, des réponses et des en-têtes pour des champs/attributs personnalisés que vous pouvez utiliser pour facetter, afficher ou trier les réponses aux FAQ. Vous pouvez également définir des champs de contrôle d'accès pour limiter la réponse à la FAQ à certains utilisateurs et groupes autorisés à voir la réponse à la FAQ.
- JSON : fichier JSON qui contient des questions, des réponses et des champs/attributs personnalisés que vous pouvez utiliser pour facetter, afficher ou trier les réponses aux FAQ. Vous pouvez également définir des champs de contrôle d'accès pour limiter la réponse à la FAQ à certains utilisateurs et groupes autorisés à voir la réponse à la FAQ.

Par exemple, ce qui suit est un fichier CSV de base qui fournit des réponses aux questions sur les cliniques gratuites à Spokane, dans l'État de Washington, aux États-Unis, et à Mountain View, dans le Missouri, aux États-Unis.

```
How many free clinics are in Spokane WA?, 13
How many free clinics are there in Mountain View Missouri?, 7
```

> **Note**
>
> Le fichier de FAQ doit être codé en UTF-8.

Rubriques

- Création de champs d'index pour un fichier FAQ
- Fichier CSV de base
- Fichier CSV personnalisé
- Fichier JSON
- Utilisation de votre fichier FAQ
- Fichiers de FAQ dans des langues autres que l'anglais

Création de champs d'index pour un fichier FAQ

> **Note**
>
> La prise en charge des fonctionnalités varie en fonction du type d'index et de l'API de recherche utilisés. Pour savoir si cette fonctionnalité est prise en charge pour le type d'index et l'API de recherche que vous utilisez, consultez la section Types d'index.

Lorsque vous utilisez un fichier CSV ou JSON personnalisé pour la saisie, vous pouvez déclarer des champs personnalisés pour vos questions de FAQ. Par exemple, vous pouvez créer un champ personnalisé qui attribue à chaque question de FAQ un département commercial. Lorsque la FAQ est renvoyée dans une réponse, vous pouvez utiliser le département comme facette pour affiner la recherche uniquement à « RH » ou « Finance », par exemple.

Un champ personnalisé doit être mappé à un champ d'index. Dans la console, vous utilisez la page de définition des facettes pour créer un champ d'index. Lorsque vous utilisez l'API, vous devez d'abord créer un champ d'index à l'aide de l'UpdateIndexAPI.

Le type de champ/attribut du fichier FAQ doit correspondre au type du champ d'index associé. Par exemple, le champ « Département » est un champ `STRING_LIST` de type. Vous devez donc fournir des valeurs pour le champ du département sous forme de liste de chaînes dans votre fichier FAQ. Vous pouvez vérifier le type des champs d'index à l'aide de la page de définition des facettes de la console ou à l'aide de l'DescribeIndexAPI.

Lorsque vous créez un champ d'index mappé à un attribut personnalisé, vous pouvez le marquer comme affichable, facetable ou triable. Vous ne pouvez pas rendre un attribut personnalisé consultable.

Outre les attributs personnalisés, vous pouvez également utiliser les champs Amazon Kendra réservés ou communs dans un fichier CSV ou JSON personnalisé. Pour plus d'informations, consultez la section Attributs ou champs du document.

Fichier CSV de base

Utilisez un fichier CSV de base lorsque vous souhaitez utiliser une structure simple pour votre FAQs. Dans un fichier CSV de base, chaque ligne comporte deux ou trois champs : une question, une réponse et un URI source facultatif qui pointe vers un document contenant plus d'informations.

Le contenu du fichier doit respecter le format commun RFC 4180 et le type MIME pour les fichiers CSV (valeurs séparées par des virgules).

Ce qui suit est un fichier de FAQ au format CSV de base.

```
How many free clinics are in Spokane WA?, 13, https://s3.region.company.com/bucket-
name/directory/faq.csv
How many free clinics are there in Mountain View Missouri?, 7, https://
s3.region.company.com/bucket-name/directory/faq.csv
```

Fichier CSV personnalisé

Utilisez un fichier CSV personnalisé lorsque vous souhaitez ajouter des champs/attributs personnalisés à vos questions de FAQ. Pour un fichier CSV personnalisé, vous utilisez une ligne d'en-tête dans votre fichier CSV pour définir les attributs supplémentaires.

Le fichier CSV doit contenir les deux champs obligatoires suivants :

- `_question`—La question fréquemment posée
- `_answer`—La réponse à la question fréquemment posée

Votre fichier CSV personnalisé peut contenir à la fois des champs Amazon Kendra réservés (sauf `_faq_id` `_data_source_id_document_title`,, et`_file_type`) et des champs personnalisés.

Voici un exemple de fichier CSV personnalisé.

```
_question,_answer,_last_updated_at,custom_string
How many free clinics are in Spokane WA?, 13, 2012-03-25T12:30:10+01:00, Note: Some
 free clinics require you to meet certain criteria in order to use their services
How many free clinics are there in Mountain View Missouri?, 7,
 2012-03-25T12:30:10+01:00, Note: Some free clinics require you to meet certain
 criteria in order to use their services
```

Le contenu du fichier personnalisé doit respecter le format commun RFC 4180 et le type MIME pour les fichiers CSV (valeurs séparées par des virgules).

La liste suivante répertorie les types de champs personnalisés :

- Date : valeurs de date et d'heure codées selon la norme ISO 8601.

  Par exemple, 2012-03-25T 12:30:10 + 01:00 est le format date-heure ISO 8601 pour le 25 mars 2012 à 12h30 (plus 10 secondes) dans le fuseau horaire d'Europe centrale.
- Long : nombres, tels que. `1234`
- Chaîne : valeurs de chaîne. Si votre chaîne contient des virgules, placez la valeur entière entre guillemets doubles («) (par exemple,`"custom attribute, and more"`).
- Liste de chaînes : liste de valeurs de chaîne. Répertoriez les valeurs dans une liste séparée par des virgules et placée entre guillemets («) (par exemple,`"item1, item2, item3"`). Si la liste ne contient qu'une seule entrée, vous pouvez omettre les guillemets (par exemple,`item1`).

Un fichier CSV personnalisé peut contenir des champs de contrôle d'accès utilisateur. Vous pouvez utiliser ces champs pour limiter l'accès à la FAQ à certains utilisateurs et groupes. Pour filtrer en fonction du contexte utilisateur, celui-ci doit fournir des informations sur l'utilisateur et le groupe dans la requête. Dans le cas contraire, tous les éléments pertinents FAQs sont renvoyés. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.

Vous trouverez ci-dessous la liste des filtres contextuels utilisateur pour FAQs :

- `_acl_user_allow`—Les utilisateurs figurant dans la liste des autorisations peuvent consulter la FAQ dans la réponse à la requête. La FAQ n'est pas renvoyée aux autres utilisateurs.
- `_acl_user_deny`—Les utilisateurs figurant dans la liste de refus ne peuvent pas voir la FAQ dans la réponse à la requête. La FAQ est renvoyée à tous les autres utilisateurs lorsqu'elle est pertinente par rapport à la requête.
- `_acl_group_allow`—Les utilisateurs membres d'un groupe autorisé peuvent consulter la FAQ dans la réponse à la requête. La FAQ n'est pas renvoyée aux utilisateurs membres d'un autre groupe.
- `_acl_group_deny`—Les utilisateurs membres d'un groupe refusé ne peuvent pas voir la FAQ dans la réponse à la requête. La FAQ est renvoyée aux autres groupes lorsqu'elle est pertinente par rapport à la requête.

Fournissez les valeurs des listes d'autorisation et de refus dans des listes séparées par des virgules et placées entre guillemets (par exemple,`"user1,user2,user3"`). Vous pouvez inclure un utilisateur ou un groupe dans une liste d'autorisation ou une liste de refus, mais pas dans les deux cas où le même utilisateur est autorisé individuellement mais également dans le cas d'un groupe refusé. Si vous incluez un utilisateur ou un groupe dans les deux, vous recevez un message d'erreur.

Voici un exemple de fichier CSV personnalisé contenant des informations contextuelles sur l'utilisateur.

```
_question, _answer, _acl_user_allow, _acl_user_deny, _acl_group_allow, _acl_group_deny
How many free clinics are in Spokane WA?, 13, "userID6201,userID7552",
 "userID1001,userID2020", groupBasicPlusRate, groupPremiumRate
```

Fichier JSON

Vous pouvez utiliser un fichier JSON pour fournir des questions, des réponses et des champs pour votre index. Vous pouvez ajouter n'importe quel champ Amazon Kendra réservé ou personnalisé à la FAQ.

Le schéma du fichier JSON est le suivant.

```
{
    "SchemaVersion": 1,
```

```
    "FaqDocuments": [
        {
            "Question": string,
            "Answer": string,
            "Attributes": {
                string: object
                additional attributes
            },
            "AccessControlList": [
               {
                   "Name": string,
                   "Type": enum( "GROUP" | "USER" ),
                   "Access": enum( "ALLOW" | "DENY" )
               },
               additional user context
            ]
        },
        additional FAQ documents
    ]
}
```

L'exemple de fichier JSON suivant montre deux documents de FAQ. L'un des documents contient uniquement la question et la réponse requises. L'autre document inclut également des informations supplémentaires sur le champ et le contexte utilisateur ou le contrôle d'accès.

```
{
    "SchemaVersion": 1,
    "FaqDocuments": [
        {
            "Question": "How many free clinics are in Spokane WA?",
            "Answer": "13"
        },
        {
            "Question": "How many free clinics are there in Mountain View Missouri?",
            "Answer": "7",
            "Attributes": {
                "_source_uri": "https://s3.region.company.com/bucket-name/directory/
faq.csv",
                "_category": "Charitable Clinics"
            },
            "AccessControlList": [
               {
                   "Name": "user@amazon.com",
```

```
                "Type": "USER",
                "Access": "ALLOW"
            },
            {
                "Name": "Admin",
                "Type": "GROUP",
                "Access": "ALLOW"
            }
        ]
    }
]
}
```

La liste suivante répertorie les types de champs personnalisés :

- Date : valeur de chaîne JSON avec des valeurs de date et d'heure codées ISO 8601. Par exemple, 2012-03-25T 12:30:10 + 01:00 est le format date-heure ISO 8601 pour le 25 mars 2012 à 12h30 (plus 10 secondes) dans le fuseau horaire d'Europe centrale.
- Long : valeur numérique JSON, telle que`1234`.
- Chaîne : valeur de chaîne JSON (par exemple,`"custom attribute"`).
- Liste de chaînes : tableau JSON de valeurs de chaîne (par exemple,`["item1,item2,item3"]`).

Un fichier JSON peut contenir des champs de contrôle d'accès utilisateur. Vous pouvez utiliser ces champs pour limiter l'accès à la FAQ à certains utilisateurs et groupes. Pour filtrer en fonction du contexte utilisateur, celui-ci doit fournir des informations sur l'utilisateur et le groupe dans la requête. Dans le cas contraire, tous les éléments pertinents FAQs sont renvoyés. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.

Vous pouvez inclure un utilisateur ou un groupe dans une liste d'autorisation ou une liste de refus, mais pas dans les deux cas où le même utilisateur est autorisé individuellement mais également dans le cas d'un groupe refusé. Si vous incluez un utilisateur ou un groupe dans les deux, vous recevez un message d'erreur.

Voici un exemple d'inclusion du contrôle d'accès utilisateur à une FAQ JSON.

```
"AccessControlList": [
                {
                    "Name": "group or user name",
                    "Type": "GROUP | USER",
```

```
                "Access": "ALLOW | DENY"
            },
            additional user context
        ]
```

Utilisation de votre fichier FAQ

Après avoir enregistré le fichier d'entrée de votre FAQ dans un compartiment S3, vous utilisez la console ou l'`CreateFaq`API pour placer les questions et réponses dans votre index. Si vous souhaitez mettre à jour une FAQ, supprimez-la et créez-la à nouveau. Vous utilisez l'`DeleteFaq`API pour supprimer une FAQ.

Vous devez fournir un IAM rôle ayant accès au compartiment S3 qui contient vos fichiers source. Vous spécifiez le rôle dans la console ou dans le `RoleArn` paramètre. Voici un exemple d'ajout d'un fichier de FAQ à un index.

Python

```
import boto3

kendra = boto3.client("kendra")

# Provide the index ID
index_id = "index-id"
# Provide the IAM role ARN required to index documents in an S3 bucket
role_arn = "arn:aws:iam::${accountId}:role/${roleName}"

# Provide the S3 bucket path information to the FAQ file
faq_path = {
    "Bucket": "bucket-name",
    "Key": "FreeClinicsUSA.csv"
}

response = kendra.create_faq(
    S3Path =  faq_path,
    Name = "FreeClinicsUSA",
    IndexId = index_id,
    RoleArn = role_arn
)

print(response)
```

Java

```
package com.amazonaws.kendra;

import software.amazon.awssdk.services.kendra.KendraClient;
import software.amazon.awssdk.services.kendra.model.CreateFaqRequest;
import software.amazon.awssdk.services.kendra.model.CreateFaqResponse;
import software.amazon.awssdk.services.kendra.model.S3Path;

public class AddFaqExample {
    public static void main(String[] args) {
        KendraClient kendra = KendraClient.builder().build();

        String indexId = "yourIndexId";
        String roleArn = "your role for accessing S3 files";

        CreateFaqRequest createFaqRequest = CreateFaqRequest
            .builder()
            .indexId(indexId)
            .name("FreeClinicsUSA")
            .roleArn(roleArn)
            .s3Path(
                S3Path
                    .builder()
                    .bucket("amzn-s3-demo-bucket")
                    .key("FreeClinicsUSA.csv")
                    .build())
            .build();

        CreateFaqResponse response = kendra.createFaq(createFaqRequest);

        System.out.println(String.format("The result of creating FAQ: %s",
 response));

    }
}
```

Fichiers de FAQ dans des langues autres que l'anglais

Vous pouvez indexer une FAQ dans une langue prise en charge. Amazon Kendra index FAQs en anglais par défaut si vous ne spécifiez aucune langue. Vous spécifiez le code de langue lorsque vous appelez l'CreateFaqopération ou vous pouvez inclure le code de langue d'une FAQ dans les

métadonnées de la FAQ sous forme de champ. Si les métadonnées d'une FAQ ne contiennent pas de code de langue spécifié dans un champ de métadonnées, la FAQ est indexée à l'aide du code de langue spécifié lorsque vous appelez l'`CreateFAQ`opération. Pour indexer un document de FAQ dans une langue prise en charge dans la console, accédez à FAQset sélectionnez Ajouter une FAQ. Vous choisissez une langue dans le menu déroulant Langue.

Création de champs de document personnalisés

> **Note**
>
> La prise en charge des fonctionnalités varie en fonction du type d'index et de l'API de recherche utilisés. Pour savoir si cette fonctionnalité est prise en charge pour le type d'index et l'API de recherche que vous utilisez, consultez la section Recherche d'index.

Vous pouvez créer des attributs ou des champs personnalisés pour vos documents dans votre index Amazon Kendra. Par exemple, vous pouvez créer un champ ou un attribut personnalisé appelé « Département » avec les valeurs « RH », « Ventes » et « Fabrication ». Si vous associez ces champs ou attributs personnalisés à votre index Amazon Kendra, vous pouvez les utiliser pour filtrer les résultats de recherche afin d'inclure des documents en fonction de l'attribut « département RH », par exemple.

Avant de pouvoir utiliser un champ ou un attribut personnalisé, vous devez d'abord créer le champ dans l'index. Utilisez la console pour modifier les mappages de champs de source de données afin d'ajouter un champ personnalisé ou utilisez l'UpdateIndexAPI pour créer le champ d'index. Vous ne pouvez pas modifier le type de données du champ une fois que vous l'avez créé.

Pour la plupart des sources de données, vous mappez les champs de la source de données externe aux champs correspondants de Amazon Kendra. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données. Pour les sources de données S3, vous pouvez créer des champs ou des attributs personnalisés à l'aide d'un fichier de métadonnées JSON.

Vous pouvez créer jusqu'à 500 champs ou attributs personnalisés.

Vous pouvez également utiliser des champs Amazon Kendra réservés ou communs. Pour plus d'informations, consultez la section Attributs ou champs du document.

Rubriques

- Mise à jour des champs de document personnalisés

Mise à jour des champs de document personnalisés

Avec l'`UpdateIndex`API, vous pouvez ajouter des champs ou des attributs personnalisés à l'aide du `DocumentMetadataConfigurationUpdates` paramètre.

L'exemple JSON suivant permet `DocumentMetadataConfigurationUpdates` d'ajouter un champ appelé « Department » à l'index.

```
"DocumentmetadataConfigurationUpdates": [
   {
       "Name": "Department",
       "Type": "STRING_VALUE"
   }
]
```

Les sections suivantes incluent des exemples d'ajout d'attributs ou de champs personnalisés à l'aide de BatchPutDocumentet pour une source de données Amazon S3.

Rubriques

- Ajouter des attributs ou des champs personnalisés avec l' BatchPutDocument API
- Ajouter des attributs ou des champs personnalisés à une source Amazon S3 de données

Ajouter des attributs ou des champs personnalisés avec l' BatchPutDocument API

Lorsque vous utilisez l'BatchPutDocumentAPI pour ajouter un document à votre index, vous spécifiez des champs ou des attributs personnalisés dans le cadre de`Attributes`. Vous pouvez ajouter plusieurs champs ou attributs lorsque vous appelez l'API. Vous pouvez créer jusqu'à 500 champs ou attributs personnalisés. L'exemple suivant est un champ ou un attribut personnalisé qui ajoute « Département » à un document.

```
"Attributes":
    {
        "Department": "HR",
        "_category": "Vacation policy"
    }
```

Ajouter des attributs ou des champs personnalisés à une source Amazon S3 de données

Lorsque vous utilisez un compartiment S3 comme source de données pour votre index, vous ajoutez des métadonnées aux documents avec des fichiers de métadonnées associés. Vous placez les fichiers JSON de métadonnées dans une structure de répertoire parallèle à celle de vos documents. Pour plus d'informations, consultez la section Métadonnées du document S3.

Vous spécifiez des champs ou des attributs personnalisés dans la structure `Attributes` JSON. Vous pouvez créer jusqu'à 500 champs ou attributs personnalisés. Par exemple, l'exemple suivant permet `Attributes` de définir trois champs ou attributs personnalisés et un champ réservé.

```
"Attributes": {
        "brand": "Amazon Basics",
        "price": 1595,
        "_category": "sports",
        "subcategories": ["outdoors", "electronics"]
    }
```

Les étapes suivantes vous expliquent comment ajouter des attributs personnalisés à une source de données Amazon S3.

Rubriques

- Étape 1 : Création d'un index Amazon Kendra
- Étape 2 : Mettre à jour l'index pour ajouter des champs de document personnalisés
- Étape 3 : créer une source de données Amazon S3 et mapper les champs de la source de données à des attributs personnalisés

Étape 1 : Création d'un index Amazon Kendra

Suivez les étapes ci-dessous Création d'un index pour créer votre index Amazon Kendra.

Étape 2 : Mettre à jour l'index pour ajouter des champs de document personnalisés

Après avoir créé un index, vous y ajoutez des champs. La procédure suivante montre comment ajouter des champs à un index à l'aide de la console et de la CLI.

Console

Pour créer des champs d'index

1. Assurez-vous d'avoir créé un index.
2. Ensuite, dans le menu de navigation de gauche, dans Gestion des données, choisissez Définition des facettes.
3. Dans le guide des paramètres des champs d'index, dans Champs d'index, choisissez Ajouter un champ pour ajouter des champs personnalisés.
4. Dans la boîte de dialogue Ajouter un champ d'index, procédez comme suit :

   - Nom du champ — Ajoutez un nom de champ.
   - Type de données : sélectionnez le type de données, qu'il s'agisse d'une chaîne, d'une liste de chaînes ou d'une date.
   - Types d'utilisation : sélectionnez les types d'utilisation, qu'ils soient facetables, consultables, affichables ou triables.

     Sélectionnez ensuite Ajouter.

   Répétez la dernière étape pour tous les autres champs que vous souhaitez mapper.

CLI

```
aws kendra update-index  \
--region $region \
--endpoint-url $endpoint \
--application-id $applicationId \
--index-id $indexId  \
--document-metadata-configuration-updates \
"[
    {
        "Name": "string",
        "Type": "STRING_VALUE"|"STRING_LIST_VALUE"|"LONG_VALUE"|"DATE_VALUE",
        "Relevance": {
            "Freshness": true|false,
            "Importance": integer,
            "Duration": "string",
            "RankOrder": "ASCENDING"|"DESCENDING",
            "ValueImportanceMap": {"string": integer
```

```
            ...}
    },
    "Search": {
        "Facetable": true|false,
        "Searchable": true|false,
        "Displayable": true|false,
        "Sortable": true|false
        }
    }
...
]"
```

Étape 3 : créer une source de données Amazon S3 et mapper les champs de la source de données à des attributs personnalisés

Pour créer une source de données Amazon S3 et y associer des champs, suivez les instructions figurant dansAmazon S3.

Si vous utilisez l'API, utilisez l'`fieldMappings`attribut ci-dessous `configuration` lorsque vous utilisez l'CreateDataSourceAPI.

Pour une vue d'ensemble de la façon dont les champs de source de données sont mappés, voirCartographie des champs de source de données.

Contrôle de l'accès des utilisateurs aux documents à l'aide de jetons

> **Note**
>
> La prise en charge des fonctionnalités varie en fonction du type d'index et de l'API de recherche utilisés. Pour savoir si cette fonctionnalité est prise en charge pour le type d'index et l'API de recherche que vous utilisez, consultez la section Types d'index.

> **Important**
>
> Les indices Amazon Kendra GenAI Enterprise Edition ne prennent pas en charge le contrôle d'accès utilisateur basé sur des jetons.

Vous pouvez contrôler quels utilisateurs ou groupes peuvent accéder à certains documents de votre index ou voir certains documents dans leurs résultats de recherche. C'est ce que l'on appelle le filtrage du contexte utilisateur. Il s'agit d'une sorte de recherche personnalisée qui a l'avantage de contrôler l'accès aux documents. Par exemple, les équipes qui recherchent des informations sur le portail de l'entreprise ne doivent pas toutes accéder aux documents top secrets de l'entreprise, et ces documents ne sont pas pertinents pour tous les utilisateurs. Seuls des utilisateurs ou des groupes d'équipes spécifiques ayant accès à des documents top secrets devraient voir ces documents dans leurs résultats de recherche.

Amazon Kendra Les indices Enterprise et Developer prennent en charge le contrôle d'accès des utilisateurs basé sur des jetons à l'aide des types de jetons suivants :

- Identifiant ouvert
- JWT avec un secret partagé
- JWT avec une clé publique
- JSON

Amazon Kendra peut être utilisé pour fournir une recherche d'entreprise sécurisée pour vos applications de récupération et de recherche. Au cours de la requête et de la récupération, Amazon Kendra filtre les résultats de recherche en fonction de la demande `AttributeFilters` et `UserContext` des informations fournies dans celle-ci. Amazon Kendra lit les listes de contrôle d'accès aux documents (ACLs) collectées par ses connecteurs lors de l'exploration et de l'ingestion. Les résultats de récupération et de recherche renvoient URLs vers les référentiels de documents originaux et de courts extraits. L'accès au document complet est toujours imposé par le référentiel d'origine.

Rubriques

- Utilisation d'OpenID
- Utilisation d'un jeton Web JSON (JWT) avec un secret partagé
- Utilisation d'un jeton Web JSON (JWT) avec une clé publique
- Utilisation de JSON

Utilisation d'OpenID

Pour configurer un Amazon Kendra index afin d'utiliser un jeton OpenID pour le contrôle d'accès, vous avez besoin de l'URL JWKS (JSON Web Key Set) du fournisseur OpenID. Dans la plupart des

cas, l'URL JWKS est au format suivant (s'ils suivent la découverte d'OpenID). `https://domain-name/.well_known/jwks.json`

Les exemples suivants montrent comment utiliser un jeton OpenID pour le contrôle d'accès des utilisateurs lorsque vous créez un index.

Console

1. Choisissez Create index pour commencer à créer un nouvel index.
2. Sur la page Spécifier les détails de l'index, donnez un nom et une description à votre index.
3. Pour IAM le rôle, sélectionnez un rôle ou sélectionnez Créer un nouveau rôle pour et spécifiez un nom de rôle pour créer un nouveau rôle. Le rôle IAM aura le préfixe « AmazonKendra - ».
4. Conservez les valeurs par défaut de tous les autres champs. Choisissez Suivant.
5. Sur la page Configurer le contrôle d'accès utilisateur, sous Paramètres de contrôle d'accès, choisissez Oui pour utiliser des jetons pour le contrôle d'accès.
6. Sous Configuration du jeton, sélectionnez OpenID comme type de jeton.
7. Spécifiez une URL de clé de signature. L'URL doit pointer vers un ensemble de clés Web JSON.
8. Facultatif sous Configuration avancée :
    a. Spécifiez un nom d'utilisateur à utiliser lors de la vérification de l'ACL.
    b. Spécifiez un ou plusieurs groupes à utiliser lors de la vérification de l'ACL.
    c. Spécifiez l'émetteur qui validera l'émetteur du jeton.
    d. Spécifiez le (s) identifiant (s) client (s). Vous devez spécifier une expression régulière correspondant à l'audience du JWT.
9. Sur la page des détails du provisionnement, choisissez Developer Edition.
10. Choisissez Créer pour créer votre index.
11. Attendez que votre index soit créé. Amazon Kendra fournit le matériel nécessaire à votre index. Cette opération peut prendre un certain temps.

CLI

Pour créer un index à l' AWS CLI aide d'un fichier d'entrée JSON, créez d'abord un fichier JSON avec les paramètres souhaités :

```
{
    "Name": "user-context",
    "Edition": "ENTERPRISE_EDITION",
    "RoleArn": "arn:aws:iam::account-id:role:/my-role",
    "UserTokenConfigurations": [
        {
            "JwtTokenTypeConfiguration": {
                "KeyLocation": "URL",
                "Issuer": "optional: specify the issuer url",
                "ClaimRegex": "optional: regex to validate claims in the token",
                "UserNameAttributeField": "optional: user",
                "GroupAttributeField": "optional: group",
                "URL": "https://example.com/.well-known/jwks.json"
            }
        }
    ],
    "UserContextPolicy": "USER_TOKEN"
}
```

Vous pouvez remplacer les noms de champs d'utilisateur et de groupe par défaut. La valeur par défaut pour `UserNameAttributeField` est « user ». La valeur par défaut pour `GroupAttributeField` est « groups ».

Ensuite, appelez `create-index` en utilisant le fichier d'entrée. Par exemple, si le nom de votre fichier JSON est`create-index-openid.json`, vous pouvez utiliser ce qui suit :

```
aws kendra create-index --cli-input-json file://create-index-openid.json
```

Python

```
response = kendra.create_index(
    Name='user-context',
    Edition='ENTERPRISE_EDITION',
    RoleArn='arn:aws:iam::account-id:role:/my-role',
    UserTokenConfigurations=[
        {
            "JwtTokenTypeConfiguration": {
                "KeyLocation": "URL",
                "Issuer": "optional: specify the issuer url",
                "ClaimRegex": "optional: regex to validate claims in the token",
                "UserNameAttributeField": "optional: user",
                "GroupAttributeField": "optional: group",
```

```
                "URL": "https://example.com/.well-known/jwks.json"
            }
        }
    ],
    UserContextPolicy='USER_TOKEN'
)
```

Utilisation d'un jeton Web JSON (JWT) avec un secret partagé

Les exemples suivants montrent comment utiliser le jeton Web JSON (JWT) avec un jeton secret partagé pour le contrôle d'accès des utilisateurs lorsque vous créez un index.

Console

1. Choisissez Create index pour commencer à créer un nouvel index.
2. Sur la page Spécifier les détails de l'index, donnez un nom et une description à votre index.
3. Pour le rôle IAM, sélectionnez un rôle ou sélectionnez Créer un nouveau rôle pour et spécifiez un nom de rôle pour créer un nouveau rôle. Le IAM rôle aura le préfixe « AmazonKendra - ».
4. Conservez les valeurs par défaut de tous les autres champs. Choisissez Suivant.
5. Sur la page Configurer le contrôle d'accès utilisateur, sous Paramètres de contrôle d'accès, choisissez Oui pour utiliser des jetons pour le contrôle d'accès.
6. Sous Configuration du jeton, sélectionnez JWT avec secret partagé comme type de jeton.
7. Sous Paramètres de signature du secret partagé, choisissez le type de secret. Vous pouvez utiliser un secret AWS Secrets Manager partagé existant ou en créer un nouveau.

   Pour créer un nouveau secret partagé, choisissez Nouveau, puis procédez comme suit :

   a. Sous Nouveau AWS Secrets Manager secret, spécifiez un nom secret. Le préfixe `AmazonKendra-` sera ajouté lorsque vous enregistrez la clé publique.

   b. Spécifiez un ID de clé. L'identifiant de clé est un indice qui indique quelle clé a été utilisée pour sécuriser la signature Web JSON du jeton.

   c. Choisissez l'algorithme de signature pour le jeton. Il s'agit de l'algorithme cryptographique utilisé pour sécuriser le jeton d'identification. Pour plus d'informations sur la norme RSA, consultez Chiffrement RSA.

d. Spécifiez un secret partagé en saisissant un secret codé en URL base64. Vous pouvez également sélectionner Générer un secret pour qu'un secret soit généré pour vous. Vous devez vous assurer que le secret est un code secret codé en URL base64.

e. (Facultatif) Spécifiez à quel moment le secret partagé est valide. Vous pouvez spécifier la date et l'heure de validité d'un secret, ou les deux. Le secret sera valide dans l'intervalle spécifié.

f. Sélectionnez Enregistrer le secret pour enregistrer le nouveau secret.

8. (Facultatif) Sous Configuration avancée :

a. Spécifiez un nom d'utilisateur à utiliser lors de la vérification de l'ACL.

b. Spécifiez un ou plusieurs groupes à utiliser lors de la vérification de l'ACL.

c. Spécifiez l'émetteur qui validera l'émetteur du jeton.

d. Spécifiez le ou les numéros de réclamation. Vous devez spécifier une expression régulière correspondant à l'audience du JWT.

9. Sur la page des détails du provisionnement, choisissez Developer Edition.

10. Choisissez Créer pour créer votre index.

11. Attendez que votre index soit créé. Amazon Kendra fournit le matériel pour votre index. Cette opération peut prendre un certain temps.

CLI

Vous pouvez utiliser un jeton JWT avec un secret partagé à l'intérieur de AWS Secrets Manager. Le secret doit être un secret codé en URL base64. Vous avez besoin de l' Secrets Manager ARN, et votre Amazon Kendra rôle doit avoir accès `GetSecretValue` à la Secrets Manager ressource. Si vous chiffrez la Secrets Manager ressource avec AWS KMS, le rôle doit également avoir accès à l'action de déchiffrement.

Pour créer un index à l' AWS CLI aide d'un fichier d'entrée JSON, créez d'abord un fichier JSON avec les paramètres souhaités :

```
{
    "Name": "user-context",
    "Edition": "ENTERPRISE_EDITION",
    "RoleArn": "arn:aws:iam::account-id:role:/my-role",
    "UserTokenConfigurations": [
        {
```

```
            "JwtTokenTypeConfiguration": {
                "KeyLocation": "SECRET_MANAGER",
                "Issuer": "optional: specify the issuer url",
                "ClaimRegex": "optional: regex to validate claims in the token",
                "UserNameAttributeField": "optional: user",
                "GroupAttributeField": "optional: group",
                "SecretManagerArn": "arn:aws:secretsmanager:us-west-2:account
 id:secret:/my-user-context-secret
            }
        }
    ],
    "UserContextPolicy": "USER_TOKEN"
}
```

Vous pouvez remplacer les noms de champs d'utilisateur et de groupe par défaut. La valeur par défaut pour `UserNameAttributeField` est « user ». La valeur par défaut pour `GroupAttributeField` est « groups ».

Ensuite, appelez `create-index` en utilisant le fichier d'entrée. Par exemple, si le nom de votre fichier JSON est`create-index-openid.json`, vous pouvez utiliser ce qui suit :

```
aws kendra create-index --cli-input-json file://create-index-openid.json
```

Le secret doit avoir le format suivant AWS Secrets Manager :

```
{
  "keys": [
    {
      "kid": "key_id",
      "alg": "HS256|HS384|HS512",
      "kty": "OCT",
      "use": "sig", //this value can be sig only for now
      "k": "secret",
      "nbf":"ISO1806 date format"
      "exp":"ISO1806 date format"
    }
  ]
}
```

Pour plus d'informations sur JWT, consultez jwt.io.

Python

Vous pouvez utiliser un jeton JWT avec un secret partagé à l'intérieur de AWS Secrets Manager. Le secret doit être un secret codé en URL base64. Vous avez besoin de l' Secrets Manager ARN, et votre Amazon Kendra rôle doit avoir accès `GetSecretValue` à la Secrets Manager ressource. Si vous chiffrez la Secrets Manager ressource avec AWS KMS, le rôle doit également avoir accès à l'action de déchiffrement.

```
response = kendra.create_index(
    Name='user-context',
    Edition='ENTERPRISE_EDITION',
    RoleArn='arn:aws:iam::account-id:role:/my-role',
    UserTokenConfigurations=[
        {
            "JwtTokenTypeConfiguration": {
                "KeyLocation": "URL",
                "Issuer": "optional: specify the issuer url",
                "ClaimRegex": "optional: regex to validate claims in the token",
                "UserNameAttributeField": "optional: user",
                "GroupAttributeField": "optional: group",
                "SecretManagerArn": "arn:aws:secretsmanager:us-west-2:account
 id:secret:/my-user-context-secret"
            }
        }
    ],
    UserContextPolicy='USER_TOKEN'
)
```

Utilisation d'un jeton Web JSON (JWT) avec une clé publique

Les exemples suivants montrent comment utiliser le jeton Web JSON (JWT) avec une clé publique pour le contrôle d'accès des utilisateurs lorsque vous créez un index. Pour plus d'informations sur JWT, consultez jwt.io.

Console

1. Choisissez Create index pour commencer à créer un nouvel index.
2. Sur la page Spécifier les détails de l'index, donnez un nom et une description à votre index.

3. Pour le rôle IAM, sélectionnez un rôle ou sélectionnez Créer un nouveau rôle pour et spécifiez un nom de rôle pour créer un nouveau rôle. Le IAM rôle aura le préfixe « AmazonKendra - ».
4. Conservez les valeurs par défaut de tous les autres champs. Choisissez Suivant.
5. Sur la page Configurer le contrôle d'accès utilisateur, sous Paramètres de contrôle d'accès, choisissez Oui pour utiliser des jetons pour le contrôle d'accès.
6. Sous Configuration du jeton, sélectionnez JWT avec clé publique comme type de jeton.
7. Sous Paramètres de signature de la clé publique, choisissez le type de secret. Vous pouvez utiliser un AWS Secrets Manager secret existant ou en créer un nouveau.

   Pour créer un nouveau secret, choisissez Nouveau, puis procédez comme suit :

   a. Sous Nouveau AWS Secrets Manager secret, spécifiez un nom secret. Le préfixe `AmazonKendra-` sera ajouté lorsque vous enregistrerez la clé publique.
   b. Spécifiez un ID de clé. L'identifiant de clé est un indice qui indique quelle clé a été utilisée pour sécuriser la signature Web JSON du jeton.
   c. Choisissez l'algorithme de signature pour le jeton. Il s'agit de l'algorithme cryptographique utilisé pour sécuriser le jeton d'identification. Pour plus d'informations sur la norme RSA, consultez Chiffrement RSA.
   d. Sous Attributs de certificat, spécifiez une chaîne de certificats facultative. La chaîne de certificats est composée d'une liste de certificats. Il commence par le certificat d'un serveur et se termine par le certificat racine.
   e. Facultatif Spécifiez l'empreinte numérique ou l'empreinte digitale. Il doit s'agir du hachage d'un certificat, calculé sur toutes les données du certificat et sur sa signature.
   f. Spécifiez l'exposant. Il s'agit de la valeur de l'exposant pour la clé publique RSA. Elle est représentée sous la forme d'une valeur codée en base64URLUInt.
   g. Spécifiez le module. Il s'agit de la valeur de l'exposant pour la clé publique RSA. Elle est représentée sous la forme d'une valeur codée en base64URLUInt.
   h. Sélectionnez Enregistrer la clé pour enregistrer la nouvelle clé.
8. Facultatif sous Configuration avancée :

   a. Spécifiez un nom d'utilisateur à utiliser lors de la vérification de l'ACL.
   b. Spécifiez un ou plusieurs groupes à utiliser lors de la vérification de l'ACL.
   c. Spécifiez l'émetteur qui validera l'émetteur du jeton.

d. Spécifiez le (s) identifiant (s) client (s). Vous devez spécifier une expression régulière correspondant à l'audience du JWT.

9. Sur la page des détails du provisionnement, choisissez Developer Edition.
10. Choisissez Créer pour créer votre index.
11. Attendez que votre index soit créé. Amazon Kendra fournit le matériel nécessaire à votre index. Cette opération peut prendre un certain temps.

CLI

Vous pouvez utiliser JWT avec une clé publique à l'intérieur d'un AWS Secrets Manager. Vous avez besoin de l' Secrets Manager ARN, et votre Amazon Kendra rôle doit avoir accès `GetSecretValue` à la Secrets Manager ressource. Si vous chiffrez la Secrets Manager ressource avec AWS KMS, le rôle doit également avoir accès à l'action de déchiffrement.

Pour créer un index à l' AWS CLI aide d'un fichier d'entrée JSON, créez d'abord un fichier JSON avec les paramètres souhaités :

```
{
    "Name": "user-context",
    "Edition": "ENTERPRISE_EDITION",
    "RoleArn": "arn:aws:iam::account id:role:/my-role",
    "UserTokenConfigurationList": [
        {
            "JwtTokenTypeConfiguration": {
                "KeyLocation": "SECRET_MANAGER",
                "Issuer": "optional: specify the issuer url",
                "ClaimRegex": "optional: regex to validate claims in the token",
                "UserNameAttributeField": "optional: user",
                "GroupAttributeField": "optional: group",
                "SecretManagerArn": "arn:aws:secretsmanager:us-west-2:account
 id:secret:/my-user-context-secret
            }
        }
    ],    "UserContextPolicy": "USER_TOKEN"
}
```

Vous pouvez remplacer les noms de champs d'utilisateur et de groupe par défaut. La valeur par défaut pour `UserNameAttributeField` est « user ». La valeur par défaut pour `GroupAttributeField` est « groups ».

Ensuite, appelez `create-index` en utilisant le fichier d'entrée. Par exemple, si le nom de votre fichier JSON est`create-index-openid.json`, vous pouvez utiliser ce qui suit :

```
aws kendra create-index --cli-input-json file://create-index-openid.json
```

Le secret doit avoir le format suivant Secrets Manager :

```
{
  "keys": [
    {
      "alg": "RS256|RS384|RS512",
      "kty": "RSA", //this can be RSA only for now
      "use": "sig", //this value can be sig only for now
      "n": "modulus of standard pem",
      "e": "exponent of standard pem",
      "kid": "key_id",
      "x5t": "certificate thumprint for x.509 cert",
      "x5c": [
        "certificate chain"
      ]
    }
  ]
}
```

Pour plus d'informations sur JWT, consultez jwt.io.

Python

```
response = kendra.create_index(
    Name='user-context',
    Edition='ENTERPRISE_EDITION',
    RoleArn='arn:aws:iam::account id:role:/my-role',
    UserTokenConfigurationList=[
        {
            "JwtTokenTypeConfiguration": {
                "KeyLocation": "URL",
                "Issuer": "optional: specify the issuer url",
                "ClaimRegex": "optional: regex to validate claims in the token",
                "UserNameAttributeField": "optional: user",
                "GroupAttributeField": "optional: group",
                "SecretManagerArn": "arn:aws:secretsmanager:us-west-2:account
 id:secret:/my-user-context-secret"
            }
```

```
        }
    ],
    UserContextPolicy='USER_TOKEN'
)
```

Utilisation de JSON

Les exemples suivants montrent comment utiliser le JSON pour le contrôle d'accès des utilisateurs lorsque vous créez un index.

> **Warning**
>
> Le jeton JSON est une charge utile non validée. Cela ne doit être utilisé que lorsque les demandes Amazon Kendra proviennent d'un serveur fiable et jamais d'un navigateur.

Console

1. Choisissez Create index pour commencer à créer un nouvel index.
2. Sur la page Spécifier les détails de l'index, donnez un nom et une description à votre index.
3. Pour IAM le rôle, sélectionnez un rôle ou sélectionnez Créer un nouveau rôle pour et spécifiez un nom de rôle pour créer un nouveau rôle. Le IAM rôle aura le préfixe « AmazonKendra - ».
4. Conservez les valeurs par défaut de tous les autres champs. Choisissez Suivant.
5. Sur la page Configurer le contrôle d'accès utilisateur, sous Paramètres de contrôle d'accès, choisissez Oui pour utiliser des jetons pour le contrôle d'accès.
6. Sous Configuration du jeton, sélectionnez JSON comme type de jeton.
7. Spécifiez un nom d'utilisateur à utiliser lors de la vérification de l'ACL.
8. Spécifiez un ou plusieurs groupes à utiliser lors de la vérification de l'ACL.
9. Choisissez Suivant.
10. Sur la page des détails du provisionnement, choisissez Developer Edition.
11. Choisissez Créer pour créer votre index.
12. Attendez que votre index soit créé. Amazon Kendra fournit le matériel nécessaire à votre index. Cette opération peut prendre un certain temps.

CLI

Pour créer un index à l' AWS CLI aide d'un fichier d'entrée JSON, créez d'abord un fichier JSON avec les paramètres souhaités :

```
{
    "Name": "user-context",
    "Edition": "ENTERPRISE_EDITION",
    "RoleArn": "arn:aws:iam::account-id:role:/my-role",
    "UserTokenConfigurations": [
        {
            "JsonTokenTypeConfiguration": {
                "UserNameAttributeField": "user",
                "GroupAttributeField": "group"
            }
        }
    ],
    "UserContextPolicy": "USER_TOKEN"
}
```

Ensuite, appelez `create-index` en utilisant le fichier d'entrée. Par exemple, si le nom de votre fichier JSON est`create-index-openid.json`, vous pouvez utiliser ce qui suit :

```
aws kendra create-index --cli-input-json file://create-index-openid.json
```

Si vous n'utilisez pas Open ID pour AWS IAM Identity Center, vous pouvez nous envoyer le jeton au format JSON. Dans ce cas, vous devez spécifier quel champ du jeton JSON contient le nom d'utilisateur et quel champ contient les groupes. Les valeurs des champs de groupe doivent être un tableau de chaînes JSON. Par exemple, si vous utilisez le protocole SAML, votre jeton sera similaire au suivant :

```
{
     "username" : "user1",
     "groups": [
        "group1",
        "group2"
     ]
}
```

Cela `TokenConfiguration` indiquerait le nom d'utilisateur et les noms des champs de groupe :

```
{
    "UserNameAttributeField":"username",
    "GroupAttributeField":"groups"
}
```

Python

```
response = kendra.create_index(
    Name='user-context',
    Edition='ENTERPRISE_EDITION',
    RoleArn='arn:aws:iam::account-id:role:/my-role',
    UserTokenConfigurations=[
        {
            "JwtTokenTypeConfiguration": {
                "UserNameAttributeField": "user",
                "GroupAttributeField": "group",
            }
        }
    ],
    UserContextPolicy='USER_TOKEN'
)
```

Création d'un connecteur de source de données

Vous pouvez créer un connecteur de source de données pour vous connecter Amazon Kendra à vos documents et les indexer. Amazon Kendra peut se connecter à Microsoft SharePoint, Google Drive et à de nombreux autres fournisseurs. Lorsque vous créez un connecteur de source de données, vous fournissez Amazon Kendra les informations de configuration requises pour vous connecter à votre référentiel source. Contrairement à l'ajout de documents directement à un index, vous pouvez régulièrement scanner la source de données pour mettre à jour l'index.

Supposons, par exemple, que vous disposiez d'un référentiel de documents fiscaux stocké dans un Amazon S3 bucket. De temps à autre, des documents existants sont modifiés et de nouveaux documents sont ajoutés au référentiel. Si vous ajoutez le référentiel en Amazon Kendra tant que source de données, vous pouvez maintenir votre index à jour en configurant des synchronisations périodiques entre votre source de données et votre index.

Vous pouvez choisir de mettre à jour un index manuellement à l'aide de la console ou de l'StartDataSourceSyncJobAPI. Sinon, vous définissez un calendrier pour mettre à jour un index et le synchroniser avec votre source de données.

Un index peut avoir plusieurs sources de données. Chaque source de données peut avoir son propre calendrier de mise à jour. Par exemple, vous pouvez mettre à jour l'index de vos documents de travail tous les jours, voire toutes les heures, tout en mettant à jour vos documents archivés manuellement chaque fois que l'archive change.

Si vous souhaitez modifier les métadonnées ou les attributs et le contenu de votre document pendant le processus d'ingestion du document, consultez la section Enrichissement Amazon Kendra personnalisé des documents.

> **Note**
>
> Chaque identifiant de document doit être unique par index. Vous ne pouvez pas créer de source de données pour indexer vos documents avec leur caractère unique, IDs puis utiliser l'`BatchPutDocument`API pour indexer les mêmes documents, ou vice versa. Vous pouvez supprimer une source de données, puis utiliser l'`BatchPutDocument`API pour indexer les mêmes documents, ou vice versa. L'utilisation du connecteur `BatchPutDocument` et `BatchDeleteDocument` APIs en combinaison avec un connecteur de source de Amazon Kendra données pour le même ensemble de documents peut entraîner des incohérences

dans vos données. Nous vous recommandons plutôt d'utiliser le connecteur de source de données Amazon Kendra personnalisé.

> **Note**
>
> Les fichiers ajoutés à l'index doivent se trouver dans un flux d'octets codé en UTF-8. Pour plus d'informations sur les documents dans Amazon Kendra, consultez la section Documents.

Définition d'un calendrier de mise à jour

Configurez votre source de données pour qu'elle soit mise à jour régulièrement avec la console ou en utilisant le `Schedule` paramètre lorsque vous créez ou mettez à jour une source de données. Le contenu du paramètre est une chaîne contenant soit une chaîne de planification `cron` au format - format, soit une chaîne vide indiquant que l'index est mis à jour à la demande. Pour le format d'une expression cron, consultez la section Expressions de planification pour les règles dans le guide de l'Amazon CloudWatch Events utilisateur. Amazon Kendra ne prend en charge que les expressions cron. Il ne prend pas en charge les expressions de taux.

Configuration d'une langue

Vous pouvez indexer tous vos documents dans une source de données dans une langue prise en charge. Vous spécifiez le code de langue pour tous vos documents dans votre source de données lorsque vous appelez CreateDataSource. Si aucun code de langue n'est spécifié dans un champ de métadonnées, le document est indexé à l'aide du code de langue spécifié pour tous les documents au niveau de la source de données. Si vous ne spécifiez aucune langue, Amazon Kendra indexe les documents dans une source de données en anglais par défaut. Pour plus d'informations sur les langues prises en charge, y compris leurs codes, voir Ajout de documents dans des langues autres que l'anglais.

Vous indexez tous vos documents dans une source de données dans une langue prise en charge à l'aide de la console. Accédez à Sources de données et modifiez votre source de données ou à Ajouter une source de données si vous ajoutez une nouvelle source de données. Sur la page Spécifier les détails de la source de données, choisissez une langue dans la liste déroulante Langue. Vous sélectionnez Mettre à jour ou continuer à saisir les informations de configuration pour vous connecter à votre source de données.

Connecteurs de source de données

Cette section explique comment vous connecter aux bases de données et Amazon Kendra aux référentiels de sources de données pris Amazon Kendra en charge à l'aide du AWS Management Console et du Amazon Kendra APIs.

Rubriques

- Schémas de modèles de sources de données
- Adobe Experience Manager
- Alfresco
- Aurora (MySQL)
- Aurora (PostgreSQL)
- Amazon FSx (Fenêtres)
- Amazon FSx (NetApp ONTAP)
- Amazon RDS/Aurora
- Amazon RDS (Microsoft SQL Server)
- Amazon RDS (MySQL)
- Amazon RDS (Oracle)
- Amazon RDS (PostgreSQL)
- Amazon S3
- Amazon Kendra Explorateur Web
- Box (Cube)
- Confluence
- Connecteur de source de données personnalisé
- Dropbox
- Drupal
- GitHub
- Gmail
- Google Drive
- IBM DB2

- Jira
- Microsoft Exchange
- Microsoft OneDrive
- Microsoft SharePoint
- Microsoft SQL Server
- Microsoft Teams
- Microsoft Yammer
- MySQL
- Oracle Database
- PostgreSQL
- Quip
- Salesforce
- ServiceNow
- Slack
- Zendesk

Schémas de modèles de sources de données

Vous trouverez ci-dessous des schémas de modèles pour les sources de données dans lesquelles les modèles sont pris en charge.

Rubriques

- Adobe Experience Managerschéma de modèle
- Amazon FSx schéma de modèle (Windows)
- Amazon FSx schéma de modèle (NetApp ONTAP)
- Alfrescoschéma de modèle
- Aurora Schéma de modèle (MySQL)
- Aurora Schéma de modèle (PostgreSQL)
- Amazon RDS Schéma de modèle (Microsoft SQL Server)
- Amazon RDS Schéma de modèle (MySQL)
- Amazon RDS schéma de modèle (Oracle)

- Amazon RDS Schéma de modèle (PostgreSQL)
- Amazon S3 schéma de modèle
- Amazon Kendra Schéma du modèle Web Crawler
- Schéma du modèle Confluence
- Schéma du modèle Dropbox
- Schéma du modèle Drupal
- GitHub schéma de modèle
- Schéma du modèle Gmail
- Schéma du modèle Google Drive
- Schéma DB2 du modèle IBM
- Schéma du modèle Microsoft Exchange
- Schéma OneDrive de modèle Microsoft
- Schéma SharePoint de modèle Microsoft
- Schéma de modèle Microsoft SQL Server
- Schéma du modèle Microsoft Teams
- Schéma du modèle Microsoft Yammer
- Schéma du modèle MySQL
- Schéma du modèle de base de données Oracle
- Schéma du modèle PostgreSQL
- Schéma du modèle Salesforce
- ServiceNow schéma de modèle
- Schéma du modèle Slack
- Schéma du modèle Zendesk

Adobe Experience Managerschéma de modèle

Vous incluez un JSON qui contient le schéma de la source de données dans le cadre de l'TemplateConfigurationobjet. Vous indiquez l'URL de l'Adobe Experience Managerhôte, le type d'authentification et indiquez si vous utilisez Adobe Experience Manager (AEM) en tant que service cloud ou AEM sur site dans le cadre de la configuration de la connexion ou des détails du point de terminaison du référentiel. Spécifiez également le type de source de donnéesAEM, un secret pour

vos informations d'authentification et les autres configurations nécessaires. Vous spécifiez ensuite `TEMPLATE` le `Type` moment de l'appel CreateDataSource.

Vous pouvez utiliser le modèle fourni dans ce guide du développeur. Pour de plus amples informations, veuillez consulter Adobe Experience ManagerSchéma JSON.

Le tableau suivant décrit les paramètres du schéma AEM JSON.

Configuration	Description
Configuration de connexion	Informations de configuration pour le point de terminaison de la source de données.
repositoryEndpointMetadata	Informations sur le point de terminaison de la source de données.
URL AEM	URL de l'Adobe Experience Managerhôte. Par exemple, si vous utilisez AEM On-Premise, vous devez inclure le nom d'hôte et le port :. https://hostname:port Ou, si vous utilisez AEM en tant que service cloud, vous pouvez utiliser l'URL de l'auteur : https://author-xxxxxx-xxxxxxx.adobeaemcloud.com.
authType	Le type d'authentification que vous utilisez, que ce soit `Basic` ou`OAuth2`.
deploymentType	Le type de Adobe Experience Manager celui que vous utilisez, `CLOUD` soit`ON_PREMISE` .
Configurations du référentiel	Informations de configuration pour le contenu de la source de données. Par exemple, configurer des types spécifiques de contenu et des mappages de champs.
• page • asset	Liste d'objets qui mappent les attributs ou les noms de champs de vos Adobe Experience Manager pages et de vos ressources aux noms de champs d' Amazon Kendra index. Pour plus

Configuration	Description
	d'informations, veuillez consulter la rubrique Mappage des champs de source de données.
Propriétés supplémentaires	Options de configuration supplémentaires pour votre contenu dans votre source de données.
timeZoneId	Si vous utilisez AEM On-Premise et que le fuseau horaire de votre serveur est différent de celui du connecteur ou de l'index Amazon Kendra AEM, vous pouvez spécifier le fuseau horaire du serveur afin de l'aligner sur le connecteur ou l'index AEM. Le fuseau horaire par défaut pour AEM On-Premise est le fuseau horaire du connecteu r ou de l' Amazon Kendra index AEM. Le fuseau horaire par défaut pour AEM en tant que service cloud est l'heure moyenne de Greenwich.
• pageRootPaths • assetRootPaths	Liste des chemins racines pour les pages et les ressources. Par exemple, le chemin racine d'une page peut être /content/sub et le chemin racine d'une ressource peut être/content/sub/ asset1.
Assets Crawlet	`true`pour ramper des actifs.
Parcourir les pages	`true`pour parcourir les pages.

Configuration	Description
• pagePathInclusionMotifs • pageNameInclusionMotifs • assetPathInclusionMotifs • assetTypeInclusionMotifs • assetNameInclusionMotifs	Liste de modèles d'expressions régulière s permettant d'inclure certaines pages et ressources dans votre source de Adobe Experience Manager données. Les pages et les ressources correspondant aux modèles sont incluses dans l'index. Les pages et les ressources qui ne correspondent pas aux modèles sont exclues de l'index. Si une page ou un actif correspond à la fois à un modèle d'inclusion et à un modèle d'exclusion, le modèle d'exclusion est prioritaire et le contenu n'est pas inclus dans l'index.
• pagePathExclusionMotifs • pageNameExclusionMotifs • assetPathExclusionMotifs • assetTypeInclusionMotifs • assetNameInclusionMotifs	Liste de modèles d'expressions régulière s permettant d'exclure certaines pages et ressources de votre source de Adobe Experience Manager données. Les pages et les ressources correspondant aux modèles sont exclues de l'index. Les pages et les ressource s qui ne correspondent pas aux modèles sont incluses dans l'index. Si une page ou un actif correspond à la fois à un modèle d'inclusion et à un modèle d'exclusion, le modèle d'exclusi on est prioritaire et le contenu n'est pas inclus dans l'index.
Composants de page	Liste des noms des composants de page spécifiques que vous souhaitez indexer.
contentFragmentVariations	Liste des noms des variantes enregistrées spécifiques des fragments de Adobe Experienc e Manager contenu que vous souhaitez indexer.

Configuration	Description
type	Type de source de données. Spécifiez `AEM` comme type de source de données.
Mode de synchronisation	Spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Vous pouvez choisir entre : • `FORCED_FULL_CRAWL` pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index. • `FULL_CRAWL` pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modificat ions de contenu et indexer le contenu modifié depuis la dernière synchronisation. • `CHANGE_LOG` pour indexer uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchroni se avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

Configuration	Description
Secrétaire Arn	Le nom de ressource Amazon (ARN) d'un AWS Secrets Manager secret qui contient les paires clé-valeur requises pour se connecter à votre Adobe Experience Manager. Pour plus d'informations sur ces paires clé-valeur, consultez les instructions de connexion pour Adobe Experience Manager.
version	Version de ce modèle actuellement prise en charge.

Adobe Experience ManagerSchéma JSON

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties":
  {
    "connectionConfiguration": {
      "type": "object",
      "properties":
      {
        "repositoryEndpointMetadata":
        {
          "type": "object",
          "properties":
          {
            "aemUrl":
            {
              "type": "string",
              "pattern": "https:.*"
            },
            "authType": {
              "type": "string",
              "enum": ["Basic", "OAuth2"]
            },
            "deploymentType": {
              "type": "string",
```

```
              "enum": ["CLOUD","ON_PREMISE"]
            }
          },
          "required":
          [
            "aemUrl",
            "authType",
            "deploymentType"
          ]
        }
      },
      "required":
      [
        "repositoryEndpointMetadata"
      ]
    },
    "repositoryConfigurations": {
      "type": "object",
      "properties":
      {
        "page":
        {
          "type": "object",
          "properties":
          {
            "fieldMappings":
            {
              "type": "array",
              "items":
              [
                {
                  "type": "object",
                  "properties":
                  {
                    "indexFieldName":
                    {
                      "type": "string"
                    },
                    "indexFieldType":
                    {
                      "type": "string",
                      "enum":
                      [
                        "STRING",
```

```
                    "STRING_LIST",
                    "DATE",
                    "LONG"
                  ]
                },
                "dataSourceFieldName":
                {
                  "type": "string"
                },
                "dateFieldFormat":
                {
                  "type": "string",
                  "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                }
              },
              "required":
              [
                "indexFieldName",
                "indexFieldType",
                "dataSourceFieldName"
              ]
            }
          ]
        }
      },
      "required":
      [
        "fieldMappings"
      ]
    },
    "asset":
    {
      "type": "object",
      "properties":
      {
        "fieldMappings":
        {
          "type": "array",
          "items":
          [
            {
              "type": "object",
              "properties":
              {
```

```
                  "indexFieldName":
                  {
                    "type": "string"
                  },
                  "indexFieldType":
                  {
                    "type": "string",
                    "enum":
                    [
                      "STRING",
                      "STRING_LIST",
                      "DATE",
                      "LONG"
                    ]
                  },
                  "dataSourceFieldName":
                  {
                    "type": "string"
                  },
                  "dateFieldFormat":
                  {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required":
                [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required":
        [
          "fieldMappings"
        ]
      }
    },
    "additionalProperties": {
      "type": "object",
```

```
      "properties":
      {
        "timeZoneId": {
          "type": "string",
          "enum": [
            "Africa/Abidjan",
            "Africa/Accra",
            "Africa/Addis_Ababa",
            "Africa/Algiers",
            "Africa/Asmara",
            "Africa/Asmera",
            "Africa/Bamako",
            "Africa/Bangui",
            "Africa/Banjul",
            "Africa/Bissau",
            "Africa/Blantyre",
            "Africa/Brazzaville",
            "Africa/Bujumbura",
            "Africa/Cairo",
            "Africa/Casablanca",
            "Africa/Ceuta",
            "Africa/Conakry",
            "Africa/Dakar",
            "Africa/Dar_es_Salaam",
            "Africa/Djibouti",
            "Africa/Douala",
            "Africa/El_Aaiun",
            "Africa/Freetown",
            "Africa/Gaborone",
            "Africa/Harare",
            "Africa/Johannesburg",
            "Africa/Juba",
            "Africa/Kampala",
            "Africa/Khartoum",
            "Africa/Kigali",
            "Africa/Kinshasa",
            "Africa/Lagos",
            "Africa/Libreville",
            "Africa/Lome",
            "Africa/Luanda",
            "Africa/Lubumbashi",
            "Africa/Lusaka",
            "Africa/Malabo",
            "Africa/Maputo",
```

```
            "Africa/Maseru",
            "Africa/Mbabane",
            "Africa/Mogadishu",
            "Africa/Monrovia",
            "Africa/Nairobi",
            "Africa/Ndjamena",
            "Africa/Niamey",
            "Africa/Nouakchott",
            "Africa/Ouagadougou",
            "Africa/Porto-Novo",
            "Africa/Sao_Tome",
            "Africa/Timbuktu",
            "Africa/Tripoli",
            "Africa/Tunis",
            "Africa/Windhoek",
            "America/Adak",
            "America/Anchorage",
            "America/Anguilla",
            "America/Antigua",
            "America/Araguaina",
            "America/Argentina/Buenos_Aires",
            "America/Argentina/Catamarca",
            "America/Argentina/ComodRivadavia",
            "America/Argentina/Cordoba",
            "America/Argentina/Jujuy",
            "America/Argentina/La_Rioja",
            "America/Argentina/Mendoza",
            "America/Argentina/Rio_Gallegos",
            "America/Argentina/Salta",
            "America/Argentina/San_Juan",
            "America/Argentina/San_Luis",
            "America/Argentina/Tucuman",
            "America/Argentina/Ushuaia",
            "America/Aruba",
            "America/Asuncion",
            "America/Atikokan",
            "America/Atka",
            "America/Bahia",
            "America/Bahia_Banderas",
            "America/Barbados",
            "America/Belem",
            "America/Belize",
            "America/Blanc-Sablon",
            "America/Boa_Vista",
```

```
            "America/Bogota",
            "America/Boise",
            "America/Buenos_Aires",
            "America/Cambridge_Bay",
            "America/Campo_Grande",
            "America/Cancun",
            "America/Caracas",
            "America/Catamarca",
            "America/Cayenne",
            "America/Cayman",
            "America/Chicago",
            "America/Chihuahua",
            "America/Ciudad_Juarez",
            "America/Coral_Harbour",
            "America/Cordoba",
            "America/Costa_Rica",
            "America/Creston",
            "America/Cuiaba",
            "America/Curacao",
            "America/Danmarkshavn",
            "America/Dawson",
            "America/Dawson_Creek",
            "America/Denver",
            "America/Detroit",
            "America/Dominica",
            "America/Edmonton",
            "America/Eirunepe",
            "America/El_Salvador",
            "America/Ensenada",
            "America/Fort_Nelson",
            "America/Fort_Wayne",
            "America/Fortaleza",
            "America/Glace_Bay",
            "America/Godthab",
            "America/Goose_Bay",
            "America/Grand_Turk",
            "America/Grenada",
            "America/Guadeloupe",
            "America/Guatemala",
            "America/Guayaquil",
            "America/Guyana",
            "America/Halifax",
            "America/Havana",
            "America/Hermosillo",
```

```
            "America/Indiana/Indianapolis",
            "America/Indiana/Knox",
            "America/Indiana/Marengo",
            "America/Indiana/Petersburg",
            "America/Indiana/Tell_City",
            "America/Indiana/Vevay",
            "America/Indiana/Vincennes",
            "America/Indiana/Winamac",
            "America/Indianapolis",
            "America/Inuvik",
            "America/Iqaluit",
            "America/Jamaica",
            "America/Jujuy",
            "America/Juneau",
            "America/Kentucky/Louisville",
            "America/Kentucky/Monticello",
            "America/Knox_IN",
            "America/Kralendijk",
            "America/La_Paz",
            "America/Lima",
            "America/Los_Angeles",
            "America/Louisville",
            "America/Lower_Princes",
            "America/Maceio",
            "America/Managua",
            "America/Manaus",
            "America/Marigot",
            "America/Martinique",
            "America/Matamoros",
            "America/Mazatlan",
            "America/Mendoza",
            "America/Menominee",
            "America/Merida",
            "America/Metlakatla",
            "America/Mexico_City",
            "America/Miquelon",
            "America/Moncton",
            "America/Monterrey",
            "America/Montevideo",
            "America/Montreal",
            "America/Montserrat",
            "America/Nassau",
            "America/New_York",
            "America/Nipigon",
```

```
            "America/Nome",
            "America/Noronha",
            "America/North_Dakota/Beulah",
            "America/North_Dakota/Center",
            "America/North_Dakota/New_Salem",
            "America/Nuuk",
            "America/Ojinaga",
            "America/Panama",
            "America/Pangnirtung",
            "America/Paramaribo",
            "America/Phoenix",
            "America/Port-au-Prince",
            "America/Port_of_Spain",
            "America/Porto_Acre",
            "America/Porto_Velho",
            "America/Puerto_Rico",
            "America/Punta_Arenas",
            "America/Rainy_River",
            "America/Rankin_Inlet",
            "America/Recife",
            "America/Regina",
            "America/Resolute",
            "America/Rio_Branco",
            "America/Rosario",
            "America/Santa_Isabel",
            "America/Santarem",
            "America/Santiago",
            "America/Santo_Domingo",
            "America/Sao_Paulo",
            "America/Scoresbysund",
            "America/Shiprock",
            "America/Sitka",
            "America/St_Barthelemy",
            "America/St_Johns",
            "America/St_Kitts",
            "America/St_Lucia",
            "America/St_Thomas",
            "America/St_Vincent",
            "America/Swift_Current",
            "America/Tegucigalpa",
            "America/Thule",
            "America/Thunder_Bay",
            "America/Tijuana",
            "America/Toronto",
```

```
            "America/Tortola",
            "America/Vancouver",
            "America/Virgin",
            "America/Whitehorse",
            "America/Winnipeg",
            "America/Yakutat",
            "America/Yellowknife",
            "Antarctica/Casey",
            "Antarctica/Davis",
            "Antarctica/DumontDUrville",
            "Antarctica/Macquarie",
            "Antarctica/Mawson",
            "Antarctica/McMurdo",
            "Antarctica/Palmer",
            "Antarctica/Rothera",
            "Antarctica/South_Pole",
            "Antarctica/Syowa",
            "Antarctica/Troll",
            "Antarctica/Vostok",
            "Arctic/Longyearbyen",
            "Asia/Aden",
            "Asia/Almaty",
            "Asia/Amman",
            "Asia/Anadyr",
            "Asia/Aqtau",
            "Asia/Aqtobe",
            "Asia/Ashgabat",
            "Asia/Ashkhabad",
            "Asia/Atyrau",
            "Asia/Baghdad",
            "Asia/Bahrain",
            "Asia/Baku",
            "Asia/Bangkok",
            "Asia/Barnaul",
            "Asia/Beirut",
            "Asia/Bishkek",
            "Asia/Brunei",
            "Asia/Calcutta",
            "Asia/Chita",
            "Asia/Choibalsan",
            "Asia/Chongqing",
            "Asia/Chungking",
            "Asia/Colombo",
            "Asia/Dacca",
```

```
            "Asia/Damascus",
            "Asia/Dhaka",
            "Asia/Dili",
            "Asia/Dubai",
            "Asia/Dushanbe",
            "Asia/Famagusta",
            "Asia/Gaza",
            "Asia/Harbin",
            "Asia/Hebron",
            "Asia/Ho_Chi_Minh",
            "Asia/Hong_Kong",
            "Asia/Hovd",
            "Asia/Irkutsk",
            "Asia/Istanbul",
            "Asia/Jakarta",
            "Asia/Jayapura",
            "Asia/Jerusalem",
            "Asia/Kabul",
            "Asia/Kamchatka",
            "Asia/Karachi",
            "Asia/Kashgar",
            "Asia/Kathmandu",
            "Asia/Katmandu",
            "Asia/Khandyga",
            "Asia/Kolkata",
            "Asia/Krasnoyarsk",
            "Asia/Kuala_Lumpur",
            "Asia/Kuching",
            "Asia/Kuwait",
            "Asia/Macao",
            "Asia/Macau",
            "Asia/Magadan",
            "Asia/Makassar",
            "Asia/Manila",
            "Asia/Muscat",
            "Asia/Nicosia",
            "Asia/Novokuznetsk",
            "Asia/Novosibirsk",
            "Asia/Omsk",
            "Asia/Oral",
            "Asia/Phnom_Penh",
            "Asia/Pontianak",
            "Asia/Pyongyang",
            "Asia/Qatar",
```

```
            "Asia/Qostanay",
            "Asia/Qyzylorda",
            "Asia/Rangoon",
            "Asia/Riyadh",
            "Asia/Saigon",
            "Asia/Sakhalin",
            "Asia/Samarkand",
            "Asia/Seoul",
            "Asia/Shanghai",
            "Asia/Singapore",
            "Asia/Srednekolymsk",
            "Asia/Taipei",
            "Asia/Tashkent",
            "Asia/Tbilisi",
            "Asia/Tehran",
            "Asia/Tel_Aviv",
            "Asia/Thimbu",
            "Asia/Thimphu",
            "Asia/Tokyo",
            "Asia/Tomsk",
            "Asia/Ujung_Pandang",
            "Asia/Ulaanbaatar",
            "Asia/Ulan_Bator",
            "Asia/Urumqi",
            "Asia/Ust-Nera",
            "Asia/Vientiane",
            "Asia/Vladivostok",
            "Asia/Yakutsk",
            "Asia/Yangon",
            "Asia/Yekaterinburg",
            "Asia/Yerevan",
            "Atlantic/Azores",
            "Atlantic/Bermuda",
            "Atlantic/Canary",
            "Atlantic/Cape_Verde",
            "Atlantic/Faeroe",
            "Atlantic/Faroe",
            "Atlantic/Jan_Mayen",
            "Atlantic/Madeira",
            "Atlantic/Reykjavik",
            "Atlantic/South_Georgia",
            "Atlantic/St_Helena",
            "Atlantic/Stanley",
            "Australia/ACT",
```

```
            "Australia/Adelaide",
            "Australia/Brisbane",
            "Australia/Broken_Hill",
            "Australia/Canberra",
            "Australia/Currie",
            "Australia/Darwin",
            "Australia/Eucla",
            "Australia/Hobart",
            "Australia/LHI",
            "Australia/Lindeman",
            "Australia/Lord_Howe",
            "Australia/Melbourne",
            "Australia/NSW",
            "Australia/North",
            "Australia/Perth",
            "Australia/Queensland",
            "Australia/South",
            "Australia/Sydney",
            "Australia/Tasmania",
            "Australia/Victoria",
            "Australia/West",
            "Australia/Yancowinna",
            "Brazil/Acre",
            "Brazil/DeNoronha",
            "Brazil/East",
            "Brazil/West",
            "CET",
            "CST6CDT",
            "Canada/Atlantic",
            "Canada/Central",
            "Canada/Eastern",
            "Canada/Mountain",
            "Canada/Newfoundland",
            "Canada/Pacific",
            "Canada/Saskatchewan",
            "Canada/Yukon",
            "Chile/Continental",
            "Chile/EasterIsland",
            "Cuba",
            "EET",
            "EST5EDT",
            "Egypt",
            "Eire",
            "Etc/GMT",
```

```
            "Etc/GMT+0",
            "Etc/GMT+1",
            "Etc/GMT+10",
            "Etc/GMT+11",
            "Etc/GMT+12",
            "Etc/GMT+2",
            "Etc/GMT+3",
            "Etc/GMT+4",
            "Etc/GMT+5",
            "Etc/GMT+6",
            "Etc/GMT+7",
            "Etc/GMT+8",
            "Etc/GMT+9",
            "Etc/GMT-0",
            "Etc/GMT-1",
            "Etc/GMT-10",
            "Etc/GMT-11",
            "Etc/GMT-12",
            "Etc/GMT-13",
            "Etc/GMT-14",
            "Etc/GMT-2",
            "Etc/GMT-3",
            "Etc/GMT-4",
            "Etc/GMT-5",
            "Etc/GMT-6",
            "Etc/GMT-7",
            "Etc/GMT-8",
            "Etc/GMT-9",
            "Etc/GMT0",
            "Etc/Greenwich",
            "Etc/UCT",
            "Etc/UTC",
            "Etc/Universal",
            "Etc/Zulu",
            "Europe/Amsterdam",
            "Europe/Andorra",
            "Europe/Astrakhan",
            "Europe/Athens",
            "Europe/Belfast",
            "Europe/Belgrade",
            "Europe/Berlin",
            "Europe/Bratislava",
            "Europe/Brussels",
            "Europe/Bucharest",
```

```
            "Europe/Budapest",
            "Europe/Busingen",
            "Europe/Chisinau",
            "Europe/Copenhagen",
            "Europe/Dublin",
            "Europe/Gibraltar",
            "Europe/Guernsey",
            "Europe/Helsinki",
            "Europe/Isle_of_Man",
            "Europe/Istanbul",
            "Europe/Jersey",
            "Europe/Kaliningrad",
            "Europe/Kiev",
            "Europe/Kirov",
            "Europe/Kyiv",
            "Europe/Lisbon",
            "Europe/Ljubljana",
            "Europe/London",
            "Europe/Luxembourg",
            "Europe/Madrid",
            "Europe/Malta",
            "Europe/Mariehamn",
            "Europe/Minsk",
            "Europe/Monaco",
            "Europe/Moscow",
            "Europe/Nicosia",
            "Europe/Oslo",
            "Europe/Paris",
            "Europe/Podgorica",
            "Europe/Prague",
            "Europe/Riga",
            "Europe/Rome",
            "Europe/Samara",
            "Europe/San_Marino",
            "Europe/Sarajevo",
            "Europe/Saratov",
            "Europe/Simferopol",
            "Europe/Skopje",
            "Europe/Sofia",
            "Europe/Stockholm",
            "Europe/Tallinn",
            "Europe/Tirane",
            "Europe/Tiraspol",
            "Europe/Ulyanovsk",
```

```
            "Europe/Uzhgorod",
            "Europe/Vaduz",
            "Europe/Vatican",
            "Europe/Vienna",
            "Europe/Vilnius",
            "Europe/Volgograd",
            "Europe/Warsaw",
            "Europe/Zagreb",
            "Europe/Zaporozhye",
            "Europe/Zurich",
            "GB",
            "GB-Eire",
            "GMT",
            "GMT0",
            "Greenwich",
            "Hongkong",
            "Iceland",
            "Indian/Antananarivo",
            "Indian/Chagos",
            "Indian/Christmas",
            "Indian/Cocos",
            "Indian/Comoro",
            "Indian/Kerguelen",
            "Indian/Mahe",
            "Indian/Maldives",
            "Indian/Mauritius",
            "Indian/Mayotte",
            "Indian/Reunion",
            "Iran",
            "Israel",
            "Jamaica",
            "Japan",
            "Kwajalein",
            "Libya",
            "MET",
            "MST7MDT",
            "Mexico/BajaNorte",
            "Mexico/BajaSur",
            "Mexico/General",
            "NZ",
            "NZ-CHAT",
            "Navajo",
            "PRC",
            "PST8PDT",
```

```
            "Pacific/Apia",
            "Pacific/Auckland",
            "Pacific/Bougainville",
            "Pacific/Chatham",
            "Pacific/Chuuk",
            "Pacific/Easter",
            "Pacific/Efate",
            "Pacific/Enderbury",
            "Pacific/Fakaofo",
            "Pacific/Fiji",
            "Pacific/Funafuti",
            "Pacific/Galapagos",
            "Pacific/Gambier",
            "Pacific/Guadalcanal",
            "Pacific/Guam",
            "Pacific/Honolulu",
            "Pacific/Johnston",
            "Pacific/Kanton",
            "Pacific/Kiritimati",
            "Pacific/Kosrae",
            "Pacific/Kwajalein",
            "Pacific/Majuro",
            "Pacific/Marquesas",
            "Pacific/Midway",
            "Pacific/Nauru",
            "Pacific/Niue",
            "Pacific/Norfolk",
            "Pacific/Noumea",
            "Pacific/Pago_Pago",
            "Pacific/Palau",
            "Pacific/Pitcairn",
            "Pacific/Pohnpei",
            "Pacific/Ponape",
            "Pacific/Port_Moresby",
            "Pacific/Rarotonga",
            "Pacific/Saipan",
            "Pacific/Samoa",
            "Pacific/Tahiti",
            "Pacific/Tarawa",
            "Pacific/Tongatapu",
            "Pacific/Truk",
            "Pacific/Wake",
            "Pacific/Wallis",
            "Pacific/Yap",
```

```
            "Poland",
            "Portugal",
            "ROK",
            "Singapore",
            "SystemV/AST4",
            "SystemV/AST4ADT",
            "SystemV/CST6",
            "SystemV/CST6CDT",
            "SystemV/EST5",
            "SystemV/EST5EDT",
            "SystemV/HST10",
            "SystemV/MST7",
            "SystemV/MST7MDT",
            "SystemV/PST8",
            "SystemV/PST8PDT",
            "SystemV/YST9",
            "SystemV/YST9YDT",
            "Turkey",
            "UCT",
            "US/Alaska",
            "US/Aleutian",
            "US/Arizona",
            "US/Central",
            "US/East-Indiana",
            "US/Eastern",
            "US/Hawaii",
            "US/Indiana-Starke",
            "US/Michigan",
            "US/Mountain",
            "US/Pacific",
            "US/Samoa",
            "UTC",
            "Universal",
            "W-SU",
            "WET",
            "Zulu",
            "EST",
            "HST",
            "MST",
            "ACT",
            "AET",
            "AGT",
            "ART",
            "AST",
```

```
            "BET",
            "BST",
            "CAT",
            "CNT",
            "CST",
            "CTT",
            "EAT",
            "ECT",
            "IET",
            "IST",
            "JST",
            "MIT",
            "NET",
            "NST",
            "PLT",
            "PNT",
            "PRT",
            "PST",
            "SST",
            "VST"
          ]
        },
        "pageRootPaths":
        {
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "assetRootPaths":
        {
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "crawlAssets":
        {
          "type": "boolean"
        },
        "crawlPages":
        {
```

```
          "type": "boolean"
        },
        "pagePathInclusionPatterns":
        {
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "pagePathExclusionPatterns":
        {
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "pageNameInclusionPatterns":
        {
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "pageNameExclusionPatterns":
        {
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "assetPathInclusionPatterns":
        {
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "assetPathExclusionPatterns":
        {
```

```
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "assetTypeInclusionPatterns":
        {
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "assetTypeExclusionPatterns":
        {
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "assetNameInclusionPatterns":
        {
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "assetNameExclusionPatterns":
        {
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "pageComponents": {
          "type": "array",
          "items": {
            "type": "object"
            }
        },
```

```
        "contentFragmentVariations": {
          "type": "array",
          "items": {
            "type": "object"
          }
        },
        "cugExemptedPrincipals": {
          "type": "array",
          "items": {
            "type": "string"
          }
        }
      },
      "required":
      []
    },
    "type": {
      "type": "string",
      "pattern": "AEM"
    },
    "enableIdentityCrawler": {
      "type": "boolean"
    },
    "syncMode": {
      "type": "string",
      "enum": [
        "FORCED_FULL_CRAWL",
        "FULL_CRAWL",
        "CHANGE_LOG"
      ]
    },
    "secretArn": {
      "type": "string",
      "minLength": 20,
      "maxLength": 2048
    }
  },
  "version": {
    "type": "string",
    "anyOf": [
      {
        "pattern": "1.0.0"
      }
    ]
```

```
  },
  "required": [
    "connectionConfiguration",
    "repositoryConfigurations",
    "syncMode",
    "additionalProperties",
    "secretArn",
    "type"
  ]
}
```

Amazon FSx schéma de modèle (Windows)

Vous incluez un JSON qui contient le schéma de la source de données dans le cadre de l'TemplateConfigurationobjet. Vous fournissez l'ID du système de fichiers dans le cadre de la configuration de la connexion ou des détails du point de terminaison du référentiel. Vous devez également spécifier le type de source de donnéesFSX, le secret de vos informations d'authentification et les autres configurations nécessaires. Vous spécifiez ensuite `TEMPLATE` le `Type` moment de l'appel CreateDataSource.

Vous pouvez utiliser le modèle fourni dans ce guide du développeur. Consultez Amazon FSx Schéma JSON (Windows).

Le tableau suivant décrit les paramètres du schéma JSON Amazon FSx (Windows).

Configuration	Description
Configuration de connexion	Informations de configuration pour le point de terminaison de la source de données.
repositoryEndpointMetadata	Informations sur le point de terminaison de la source de données.
fileSystemId	Identifiant du système de Amazon FSx fichiers. Vous trouverez l'ID de votre système de fichiers sur le tableau de bord des systèmes de fichiers de la Amazon FSx console.

Configuration	Description
fileSystemType	Type de système de Amazon FSx fichiers. À utiliser Windows File Server comme type de système de fichiers, spécifiez`WINDOWS`.
Configurations du référentiel	Informations de configuration pour le contenu de la source de données. Par exemple, configurer des types spécifiques de contenu et des mappages de champs.
Tous	Liste d'objets qui mappent les attributs ou les noms de champs de vos fichiers dans votre source de Amazon FSx données aux noms de champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.
Propriétés supplémentaires	Options de configuration supplémentaires pour votre contenu dans votre source de données.
isCrawlAcl	`true`pour analyser les informations de la liste de contrôle d'accès (ACL) de vos documents , si vous disposez d'une ACL et que vous souhaitez l'utiliser pour le contrôle d'accès. L'ACL indique les documents auxquels les utilisateurs et les groupes peuvent accéder. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.

Configuration	Description
Modèles d'inclusion	Liste de modèles d'expressions régulières permettant d'inclure certains fichiers dans votre source de Amazon FSx données. Les fichiers qui correspondent aux modèles sont inclus dans l'index. Les fichiers qui ne correspon dent pas aux modèles sont exclus de l'index. Si un fichier correspond à la fois à un modèle d'inclusion et d'exclusion, le modèle d'exclusi on a la priorité et le fichier n'est pas inclus dans l'index.
Schémas d'exclusion	Liste de modèles d'expressions régulières permettant d'exclure certains fichiers de votre source de Amazon FSx données. Les fichiers qui correspondent aux modèles sont exclus de l'index. Les fichiers qui ne correspondent pas aux modèles sont inclus dans l'index. Si un fichier correspond à la fois à un modèle d'exclusion et à un modèle d'inclusion, le modèle d'exclusion est prioritaire et le fichier n'est pas inclus dans l'index.
enableIdentityCrawler	`true`utiliser Amazon Kendra le moteur de recherche d'identité pour synchroniser les identity/principal informations relatives aux utilisateurs et aux groupes ayant accès à certains documents. Si le robot d'identification est désactivé, tous les documents peuvent être consultés publiquement. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'identification est désactivé , vous pouvez également utiliser l'PutPrinci palMappingAPI pour télécharger les informati ons d'accès des utilisateurs et des groupes.

Configuration	Description
Mode de synchronisation	Spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Vous pouvez choisir entre : • `FORCED_FULL_CRAWL` pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index. • `FULL_CRAWL` pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
type	Type de source de données. Pour les sources de données du système de fichiers Windows, spécifiez`FSX`.

Amazon FSx Schéma JSON (Windows)

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "fileSystemId": {
```

```
            "type": "string",
            "pattern": "fs-.*"
          },
          "fileSystemType": {
            "type": "string",
            "pattern": "WINDOWS"
          }
        },
        "required": ["fileSystemId", "fileSystemType"]
      }
    }
  },
  "repositoryConfigurations": {
    "type": "object",
    "properties": {
      "All": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string",
                    "enum": ["STRING", "STRING_LIST", "DATE"]
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  },
                  "dateFieldFormat": {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required": [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
```

```
              }
            ]
          }
        },
        "required": ["fieldMappings"]
      }
    },
    "required": ["All"]
  },
  "additionalProperties": {
    "type": "object",
    "properties": {
      "isCrawlAcl": {
        "type": "boolean"
      },
      "exclusionPatterns": {
        "type": "array",
        "items": {
          "type": "string"
        }
      },
      "inclusionPatterns": {
        "type": "array",
        "items": {
          "type": "string"
        }
      }
    },
    "required": []
  },
  "enableIdentityCrawler": {
    "type": "boolean"
  },
  "syncMode": {
    "type": "string",
    "enum": [
      "FORCED_FULL_CRAWL",
      "FULL_CRAWL"
    ]
  },
  "type" : {
    "type" : "string",
    "pattern": "FSX"
  }
```

```
  },
  "version": {
    "type": "string",
    "anyOf": [
      {
        "pattern": "1.0.0"
      }
    ]
  },
  "required": [
    "connectionConfiguration",
    "repositoryConfigurations",
    "syncMode",
    "enableIdentityCrawler",
    "additionalProperties",
    "type"
  ]
}
```

Amazon FSx schéma de modèle (NetApp ONTAP)

Vous incluez un JSON qui contient le schéma de la source de données dans le cadre de l'TemplateConfigurationobjet. Vous fournissez l'ID du système de fichiers et la machine virtuelle de stockage (SVM) dans le cadre de la configuration de la connexion ou des détails du point de terminaison du référentiel. Vous devez également spécifier le type de source de donnéesFSXONTAP, le secret de vos informations d'authentification et les autres configurations nécessaires. Vous spécifiez ensuite TEMPLATE le Type moment de l'appel CreateDataSource.

Vous pouvez utiliser le modèle fourni dans ce guide du développeur. Consultez Amazon FSx Schéma JSON (NetApp ONTAP).

Le tableau suivant décrit les paramètres du schéma JSON Amazon FSx (NetApp ONTAP).

Configuration	Description
Configuration de connexion	Informations de configuration pour le point de terminaison de la source de données.
repositoryEndpointMetadata	Informations sur le point de terminaison de la source de données.

Configuration	Description
fileSystemId	Identifiant du système de Amazon FSx fichiers. Vous trouverez l'ID de votre système de fichiers sur le tableau de bord des systèmes de fichiers de la Amazon FSx console. Pour plus d'informations sur la création d'un système de fichiers dans la Amazon FSx console pour NetApp ONTAP, consultez le guide de démarrage d' NetAppONTAP dans le guide de l'FSx for ONTAP utilisateur.
fileSystemType	Type de système de Amazon FSx fichiers. À utiliser NetApp ONTAP comme type de système de fichiers, spécifiez`ONTAP`.
SVMID	Identifiant de la machine virtuelle de stockage (SVM) utilisé avec votre système de Amazon FSx fichiers pourNetApp ONTAP. Vous pouvez trouver votre identifiant de SVM en accédant au tableau de bord des systèmes de fichiers de la Amazon FSx console, en sélectionnant l'identifiant de votre système de fichiers, puis en sélectionnant Machines virtuelles de stockage. Pour plus d'informations sur la création d'un système de fichiers dans la Amazon FSx console pourNetApp ONTAP, consultez le guide de démarrage d' NetAppONTAP dans le guide de l'FSx for ONTAP utilisateur.
Type de protocole	Que vous utilisiez le protocole CIFS (Common Internet File System) pour Windows ou le protocole NFS (Network File System) pour Linux.

Configuration	Description
Configurations du référentiel	Informations de configuration pour le contenu de la source de données. Par exemple, configurer des types spécifiques de contenu et des mappages de champs.
dans le fichier	Liste d'objets qui mappent les attributs ou les noms de champs de vos fichiers dans votre source de Amazon FSx données aux noms de champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données. Les noms des champs de source de données doivent figurer dans les métadonnées personnalisées de vos fichiers.
Propriétés supplémentaires	Options de configuration supplémentaires pour votre contenu dans votre source de données.
CrawlacL	`true`pour analyser les informations de la liste de contrôle d'accès (ACL) de vos documents , si vous disposez d'une ACL et que vous souhaitez l'utiliser pour le contrôle d'accès. L'ACL indique les documents auxquels les utilisateurs et les groupes peuvent accéder. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.

Configuration	Description
Modèles d'inclusion	Liste de modèles d'expressions régulières permettant d'inclure certains fichiers dans votre source de Amazon FSx données. Les fichiers qui correspondent aux modèles sont inclus dans l'index. Les fichiers qui ne correspon dent pas aux modèles sont exclus de l'index. Si un fichier correspond à la fois à un modèle d'inclusion et d'exclusion, le modèle d'exclusi on a la priorité et le fichier n'est pas inclus dans l'index.
Schémas d'exclusion	Liste de modèles d'expressions régulières permettant d'exclure certains fichiers de votre source de Amazon FSx données. Les fichiers qui correspondent aux modèles sont exclus de l'index. Les fichiers qui ne correspondent pas aux modèles sont inclus dans l'index. Si un fichier correspond à la fois à un modèle d'exclusion et à un modèle d'inclusion, le modèle d'exclusion est prioritaire et le fichier n'est pas inclus dans l'index.
type	Type de source de données. Pour les sources de données du système de NetApp ONTAP fichiers, spécifiez`FSXONTAP`.

Configuration	Description
Mode de synchronisation	Spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Vous pouvez choisir entre : • `FORCED_FULL_CRAWL` pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index. • `FULL_CRAWL` pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modificat ions de contenu et indexer le contenu modifié depuis la dernière synchronisation.

Configuration	Description

Secrétaire Arn	Le nom de ressource Amazon (ARN) d'un AWS Secrets Manager secret qui contient les paires clé-valeur requises pour se connecter à votre système de Amazon FSx fichiers. Le secret doit contenir une structure JSON avec les clés suivantes :
	`{ "username": " user@corp.example.com ", "password": " password" }`
	Si vous utilisez le protocole NFS pour votre système de Amazon FSx fichiers, le secret est stocké dans une structure JSON avec les clés suivantes :
	`{ "leftId": "left ID", "rightId": " right ID", "preSharedKey": " pre-shared key " }`

Amazon FSx Schéma JSON (NetApp ONTAP)

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "fileSystemId": {
```

```
            "type": "string",
              "pattern": "^(fs-[0-9a-f]{8,21})$"
          },
          "fileSystemType": {
            "type": "string",
            "enum": ["ONTAP"]
          },
          "svmId": {
            "type": "string",
            "pattern": "^(svm-[0-9a-f]{17,21})$"
          },
          "protocolType": {
            "type": "string",
            "enum": [
              "CIFS",
              "NFS"
            ]
          }
        },
        "required": [
          "fileSystemId",
          "fileSystemType"
        ]
      }
    },
    "required": [
      "repositoryEndpointMetadata"
    ]
  },
  "repositoryConfigurations": {
    "type": "object",
    "properties": {
      "file": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string",
                    "pattern": "^([a-zA-Z_]{1,20})$"
```

```
                    },
                    "indexFieldType": {
                      "type": "string",
                      "enum": [
                        "STRING",
                        "STRING_LIST",
                        "DATE",
                        "LONG"
                      ]
                    },
                    "dataSourceFieldName": {
                      "type": "string",
                      "pattern": "^([a-zA-Z_]{1,20})$"
                    },
                    "dateFieldFormat": {
                      "type": "string",
                      "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                    }
                  },
                  "required": [
                    "indexFieldName",
                    "indexFieldType",
                    "dataSourceFieldName"
                  ]
                }
              ],
              "maxItems": 50
            }
          },
          "required": [
            "fieldMappings"
          ]
        }
      },
      "required": [
        "file"
      ]
    },
    "additionalProperties": {
      "type": "object",
      "properties": {
        "crawlAcl": {
          "type": "boolean"
        },
```

```
        "inclusionPatterns": {
          "type": "array",
          "items": {
            "type": "string",
            "maxLength": 30
          },
          "maxItems": 100
        },
        "exclusionPatterns": {
          "type": "array",
          "items": {
            "type": "string",
            "maxLength": 30
          },
          "maxItems": 100
        }
      }
    },
    "type": {
      "type": "string",
      "pattern": "FSXONTAP"
    },
    "syncMode": {
      "type": "string",
      "enum": [
        "FORCED_FULL_CRAWL",
        "FULL_CRAWL"
      ]
    },
    "secretArn": {
      "type": "string",
      "pattern": "arn:aws:secretsmanager:.*"
    }
  },
  "required": [
    "connectionConfiguration",
    "repositoryConfigurations",
    "additionalProperties",
    "secretArn",
    "type"
  ]
}
```

Alfrescoschéma de modèle

Vous incluez un JSON qui contient le schéma de la source de données dans le cadre de l'TemplateConfigurationobjet. Vous indiquez l'ID du Alfresco site, l'URL du référentiel, l'URL de l'interface utilisateur, le type d'authentification, si vous utilisez le cloud ou sur site, et le type de contenu que vous souhaitez analyser. Vous le fournissez dans le cadre de la configuration de la connexion ou des détails du point de terminaison du référentiel. Spécifiez également le type de source de donnéesALFRESCO, un secret pour vos informations d'authentification et les autres configurations nécessaires. Vous spécifiez ensuite TEMPLATE le Type moment de l'appel CreateDataSource.

Vous pouvez utiliser le modèle fourni dans ce guide du développeur. Consultez AlfrescoSchéma JSON.

Le tableau suivant décrit les paramètres du schéma Alfresco JSON.

Configuration	Description
Configuration de connexion	Informations de configuration pour le point de terminaison de la source de données.
repositoryEndpointMetadata	Informations sur le point de terminaison de la source de données.
ID du site	L'identifiant du site Alfresco.
URL de retour	URL de votre Alfresco dépôt. Vous pouvez obtenir l'URL du dépôt auprès de votre Alfresco administrateur. Par exemple, si vous utilisez Alfresco le Cloud (PaaS), l'URL du référentiel peut être https://company.alfrescocloud.com. Ou, si vous utilisez Alfresco On-Premises, l'URL du référentiel peut être https://company-alfresco-instance.company-domain.suffix:port.
webAppUrl	URL de votre interface Alfresco utilisateur. Vous pouvez obtenir l'URL de Alfresco l'interface utilisateur auprès de votre Alfresco administr

Configuration	Description
	ateur. Par exemple, l'URL de l'interface utilisate ur peut être https://example.com.
repositoryAdditionalProperties	Propriétés supplémentaires pour se connecter au point de terminaison repository/data source.
authType	Le type d'authentification que vous utilisez, que ce soit `OAuth2` ou`Basic`.
type (déploiement)	Le type de Alfresco celui que vous utilisez, que ce soit `PAAS` ou`ON-PREM`.
Type de rampe	Le type de contenu que vous souhaitez explorer, que ce soit `ASPECT` (contenu marqué d'un « Aspects »Alfresco), `SITE_ID` (contenu d'un Alfresco site spécifique) ou `ALL_SITES` (contenu de tous vos Alfresco sites).
Configurations du référentiel	Informations de configuration pour le contenu de la source de données. Par exemple, configurer des types spécifiques de contenu et des mappages de champs.
• document • comment	Liste d'objets qui mappent les attributs ou les noms de champs de vos documents et commentaires Alfresco aux noms de champs d' Amazon Kendra index. Pour plus d'informa tions, veuillez consulter la rubrique Mappage des champs de source de données.
Propriétés supplémentaires	Options de configuration supplémentaires pour votre contenu dans votre source de données.
Nom de l'aspect	Le nom d'un « Aspect » spécifique que vous souhaitez indexer.

Configuration	Description
Propriétés de l'aspect	Liste des propriétés de contenu « Aspect » spécifiques que vous souhaitez indexer.
enableFineGrainedContrôle	`true`pour explorer « Aspects ».
isCrawlComment	`true`pour explorer les commentaires.
• inclusionFileNameMotifs • inclusionFileTypeMotifs • inclusionFilePathMotifs	Liste de modèles d'expressions régulières permettant d'inclure certains fichiers dans votre source de Alfresco données. Les fichiers qui correspondent aux modèles sont inclus dans l'index. Les fichiers qui ne correspondent pas aux modèles sont exclus de l'index. Si un fichier correspond à la fois à un modèle d'inclusion et à un modèle d'exclusion, le modèle d'exclusion est prioritaire et le fichier n'est pas inclus dans l'index.
• exclusionFileNameMotifs • exclusionFileTypeMotifs • exclusionFilePathMotifs	Liste de modèles d'expressions régulières permettant d'exclure certains fichiers de votre source de Alfresco données. Les fichiers qui correspondent aux modèles sont exclus de l'index. Les fichiers qui ne correspondent pas aux modèles sont inclus dans l'index. Si un fichier correspond à la fois à un modèle d'inclusion et à un modèle d'exclusion, le modèle d'exclusion est prioritaire et le fichier n'est pas inclus dans l'index.
type	Type de source de données. Spécifiez `ALFRESCO` comme type de source de données.

Configuration	Description

Secrétaire Arn	Le nom de ressource Amazon (ARN) d'un AWS Secrets Manager secret qui contient les paires clé-valeur requises pour se connecter à votre. Alfresco Le secret doit contenir une structure JSON avec les clés suivantes :
	
	Si vous utilisez l'authentification de base :
	
	`{ "username": " user name", "password": " password" }`
	
	Si vous utilisez l'authentification OAuth 2.0 :
	
	`{ "clientId": " client ID", "clientSecret": " client secret", "tokenUrl": " token URL" }`

Configuration	Description
Mode de synchronisation	Spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Vous pouvez choisir entre : • `FORCED_FULL_CRAWL` pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index. • `FULL_CRAWL` pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modificat ions de contenu et indexer le contenu modifié depuis la dernière synchronisation.
enableIdentityCrawler	`true`utiliser Amazon Kendra le moteur de recherche d'identité pour synchroniser les identity/principal informations relatives aux utilisateurs et aux groupes ayant accès à certains documents. Si le robot d'identification est désactivé, tous les documents peuvent être consultés publiquement. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'identification est désactivé , vous pouvez également utiliser l'PutPrinci palMappingAPI pour télécharger les informati ons d'accès des utilisateurs et des groupes.
version	Version de ce modèle actuellement prise en charge.

AlfrescoSchéma JSON

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "siteId": {
              "type": "string"
            },
            "repoUrl": {
              "type": "string"
            },
            "webAppUrl": {
              "type": "string"
            },
            "repositoryAdditionalProperties": {
              "type": "object",
              "properties": {
                "authType": {
                  "type": "string",
                  "enum": [
                    "OAuth2",
                    "Basic"
                  ]
                },
                "type": {
                  "type": "string",
                  "enum": [
                    "PAAS",
                    "ON_PREM"
                  ]
                },
                "crawlType": {
                  "type": "string",
                  "enum": [
                    "ASPECT",
                    "SITE_ID",
                    "ALL_SITES"
```

```
                ]
              }
            }
          }
        }
      }
    },
    "required": [
      "repositoryEndpointMetadata"
    ]
  },
  "repositoryConfigurations": {
    "type": "object",
    "properties": {
      "document": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": {
              "anyOf": [
                {
                  "type": "object",
                  "properties": {
                    "indexFieldName": {
                      "type": "string"
                    },
                    "indexFieldType": {
                      "type": "string",
                      "enum": [
                        "STRING",
                        "DATE",
                        "STRING_LIST",
                        "LONG"
                      ]
                    },
                    "dataSourceFieldName": {
                      "type": "string"
                    },
                    "dateFieldFormat": {
                      "type": "string",
                      "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                    }
                  },
```

```
                "required": [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        }
      },
      "required": [
        "fieldMappings"
      ]
    },
    "comment": {
      "type": "object",
      "properties": {
        "fieldMappings": {
          "type": "array",
          "items": {
            "anyOf": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string",
                    "enum": [
                      "STRING",
                      "DATE",
                      "STRING_LIST",
                      "LONG"
                    ]
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  },
                  "dateFieldFormat": {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
```

```
                  "required": [
                    "indexFieldName",
                    "indexFieldType",
                    "dataSourceFieldName"
                  ]
                }
              ]
            }
          }
        },
        "required": [
          "fieldMappings"
        ]
      }
    }
  },
  "additionalProperties": {
    "type": "object",
    "properties": {
      "aspectName": {
        "type": "string"
      },
      "aspectProperties": {
        "type": "array"
      },
      "enableFineGrainedControl": {
        "type": "boolean"
      },
      "isCrawlComment": {
        "type": "boolean"
      },
      "inclusionFileNamePatterns": {
        "type": "array"
      },
      "exclusionFileNamePatterns": {
        "type": "array"
      },
      "inclusionFileTypePatterns": {
        "type": "array"
      },
      "exclusionFileTypePatterns": {
        "type": "array"
      },
      "inclusionFilePathPatterns": {
```

```
          "type": "array"
        },
        "exclusionFilePathPatterns": {
          "type": "array"
        }
      }
    },
    "type": {
      "type": "string",
      "pattern": "ALFRESCO"
    },
    "secretArn": {
      "type": "string",
      "minLength": 20,
      "maxLength": 2048
    },
    "syncMode": {
      "type": "string",
      "enum": [
        "FORCED_FULL_CRAWL",
        "FULL_CRAWL"
      ]
    },
    "enableIdentityCrawler": {
      "type": "boolean"
    },
    "version": {
      "type": "string",
      "anyOf": [
        {
          "pattern": "1.0.0"
        }
      ]
    }
  },
  "required": [
    "connectionConfiguration",
    "repositoryConfigurations",
    "additionalProperties",
    "type",
    "secretArn"
  ]
}
```

Aurora Schéma de modèle (MySQL)

Vous incluez un JSON qui contient le schéma de la source de données dans le cadre de l'TemplateConfigurationobjet. Spécifiez le type de source de donnéesJDBC, le type de base de donnéesmysql, un secret pour vos informations d'authentification et les autres configurations nécessaires. Vous spécifiez ensuite TEMPLATE le Type moment de l'appel CreateDataSource.

Vous pouvez utiliser le modèle fourni dans ce guide du développeur. Consultez Aurora Schéma JSON (MySQL).

Le tableau suivant décrit les paramètres du schéma JSON Aurora (MySQL).

Configuration	Description
Configuration de connexion	Informations de configuration pour le point de terminaison de la source de données.
repositoryEndpointMetadata	Informations de configuration requises pour connecter votre source de données. • DBType : le type de base de données Java que vous utilisez, que ce soit, mysqldb2, postgresql ou. oracle sqlserver • DBHost : nom d'hôte de la base de données. • DBPort : port de base de données. • DBInstance : instance de base de données.
Configurations du référentiel	Informations de configuration pour le contenu de la source de données. Par exemple, configurer des types spécifiques de contenu et des mappages de champs. Spécifiez le type de source de données et l'ARN secret.
document	Liste d'objets qui mappent les attributs ou les noms de champs du contenu de votre base de données aux noms de champs d' Amazon Kendra index. Pour plus d'informations, veuillez

Configuration	Description
	consulter la rubrique Mappage des champs de source de données.
Propriétés supplémentaires	Options de configuration supplémentaires pour votre contenu dans votre source de données. À utiliser pour inclure ou exclure un contenu spécifique dans la source de données de votre base de données.
Clé primaire	Fournissez la clé primaire pour la table de base de données. Cela identifie une table au sein de votre base de données.
Colonne de titre	Indiquez le nom de la colonne du titre du document dans votre table de base de données.
Colonne Body	Indiquez le nom de la colonne du titre du document dans votre table de base de données.
sqlQuery	Entrez des instructions de requête SQL telles que les opérations SELECT et JOIN. La taille des requêtes SQL doit être inférieure à 32 Ko. Amazon Kendra analysera tout le contenu de la base de données correspondant à votre requête.
Colonne d'horodatage	Entrez le nom de la colonne contenant les horodatages. Amazon Kendra utilise les informations d'horodatage pour détecter les modifications apportées à votre contenu et synchroniser uniquement le contenu modifié.

Configuration	Description
Format d'horodatage	Entrez le nom de la colonne contenant les formats d'horodatage à utiliser pour détecter les modifications de contenu et resynchroniser votre contenu.
timezone	Entrez le nom de la colonne contenant les fuseaux horaires du contenu à analyser.
changeDetectingColumns	Entrez le nom des colonnes qui Amazon Kendra seront utilisées pour détecter les modifications de contenu. Amazon Kendra réindexera le contenu en cas de modification de l'une de ces colonnes
allowedUsersColumns	Entrez le nom de la colonne contenant l'utilisa teur devant IDs être autorisé à accéder au contenu.
allowedGroupsColumn	Entrez le nom de la colonne contenant l'utilisa teur devant IDs être autorisé à accéder au contenu.
source URIColumn	Entrez le nom de la colonne contenant la source URLs à indexer.
isSslEnabled	Entrez des instructions de requête SQL telles que les opérations SELECT et JOIN. La taille des requêtes SQL doit être inférieure à 32 Ko. Amazon Kendra analysera tout le contenu de la base de données correspondant à votre requête.
type	Type de source de données. Spécifiez `JDBC` comme type de source de données.

Configuration	Description
Mode de synchronisation	Spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Vous pouvez choisir entre : • `FORCED_FULL_CRAWL` pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index. • `FULL_CRAWL` pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modificat ions de contenu et indexer le contenu modifié depuis la dernière synchronisation. • `CHANGE_LOG` pour indexer uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchroni se avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

Configuration	Description
Secrétaire Arn	Le nom de ressource Amazon (ARN) d'un secret Secrets Manager contenant le nom d'utilisateur et le mot de passe requis pour se connecter à votre base de données. Le secret doit contenir une structure JSON avec les clés suivantes : `{ "user name": "database user name", "password": " password" }`
version	Version du modèle actuellement prise en charge.

Aurora Schéma JSON (MySQL)

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "dbType": {
              "type": "string",
              "enum": [
                "mysql",
                "db2",
                "postgresql",
                "oracle",
                "sqlserver"
              ]
            },
            "dbHost": {
```

```
            "type": "string"
          },
          "dbPort": {
            "type": "string"
          },
          "dbInstance": {
            "type": "string"
          }
        },
        "required": [
          "dbType",
          "dbHost",
          "dbPort",
          "dbInstance"
        ]
      }
    },
    "required": [
      "repositoryEndpointMetadata"
    ]
  },
  "repositoryConfigurations": {
    "type": "object",
    "properties": {
      "document": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string"
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  }
                },
                "required": [
```

```
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      }
    },
    "required": [
    ]
  },
  "additionalProperties": {
    "type": "object",
    "properties": {
      "primaryKey": {
        "type": "string"
      },
      "titleColumn": {
        "type": "string"
      },
      "bodyColumn": {
        "type": "string"
      },
      "sqlQuery": {
        "type": "string",
        "not": {
          "pattern": ";+"
        }
      },
      "timestampColumn": {
        "type": "string"
      },
      "timestampFormat": {
        "type": "string"
      },
      "timezone": {
        "type": "string"
      },
      "changeDetectingColumns": {
```

```
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "allowedUsersColumn": {
          "type": "string"
        },
        "allowedGroupsColumn": {
          "type": "string"
        },
        "sourceURIColumn": {
          "type": "string"
        },
        "isSslEnabled": {
          "type": "boolean"
        }
      },
      "required": ["primaryKey", "titleColumn", "bodyColumn", "sqlQuery"]
    },
    "type" : {
      "type" : "string",
      "pattern": "JDBC"
    },
    "syncMode": {
      "type": "string",
      "enum": [
        "FORCED_FULL_CRAWL",
        "FULL_CRAWL",
        "CHANGE_LOG"
      ]
    },
    "secretArn": {
      "type": "string"
    }
  },
  "version": {
    "type": "string",
    "anyOf": [
      {
        "pattern": "1.0.0"
      }
    ]
  },
```

```
  "required": [
      "connectionConfiguration",
      "repositoryConfigurations",
      "syncMode",
      "additionalProperties",
      "secretArn",
      "type"
  ]
}
```

Aurora Schéma de modèle (PostgreSQL)

Vous incluez un JSON qui contient le schéma de la source de données dans le cadre de l'TemplateConfigurationobjet. Spécifiez le type de source de donnéesJDBC, le type de base de donnéespostgresql, un secret pour vos informations d'authentification et les autres configurations nécessaires. Vous spécifiez ensuite TEMPLATE le Type moment de l'appel CreateDataSource.

Vous pouvez utiliser le modèle fourni dans ce guide du développeur. Consultez Aurora Schéma JSON (PostgreSQL).

Le tableau suivant décrit les paramètres du schéma Aurora JSON (PostgreSQL).

Configuration	Description
Configuration de connexion	Informations de configuration pour le point de terminaison de la source de données.
repositoryEndpointMetadata	Informations de configuration requises pour connecter votre source de données. • DBType : le type de base de données Java que vous utilisez, que ce soit, mysqldb2, postgresql ou. oracle sqlserver • DBHost : nom d'hôte de la base de données. • DBPort : port de base de données. • DBInstance : instance de base de données.
Configurations du référentiel	Informations de configuration pour le contenu de la source de données. Par exemple,

Configuration	Description
	configurer des types spécifiques de contenu et des mappages de champs. Spécifiez le type de source de données et l'ARN secret.
document	Liste d'objets qui mappent les attributs ou les noms de champs du contenu de votre base de données aux noms de champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.
Propriétés supplémentaires	Options de configuration supplémentaires pour votre contenu dans votre source de données. À utiliser pour inclure ou exclure un contenu spécifique dans la source de données de votre base de données.
Clé primaire	Fournissez la clé primaire pour la table de base de données. Cela identifie une table au sein de votre base de données.
Colonne de titre	Indiquez le nom de la colonne du titre du document dans votre table de base de données.
Colonne Body	Indiquez le nom de la colonne du titre du document dans votre table de base de données.
sqlQuery	Entrez des instructions de requête SQL telles que les opérations SELECT et JOIN. La taille des requêtes SQL doit être inférieure à 32 Ko. Amazon Kendra analysera tout le contenu de la base de données correspondant à votre requête.

Configuration	Description
Colonne d'horodatage	Entrez le nom de la colonne contenant les horodatages. Amazon Kendra utilise les informations d'horodatage pour détecter les modifications apportées à votre contenu et synchroniser uniquement le contenu modifié.
Format d'horodatage	Entrez le nom de la colonne contenant les formats d'horodatage à utiliser pour détecter les modifications de contenu et resynchroniser votre contenu.
timezone	Entrez le nom de la colonne contenant les fuseaux horaires du contenu à analyser.
changeDetectingColumns	Entrez le nom des colonnes qui Amazon Kendra seront utilisées pour détecter les modifications de contenu. Amazon Kendra réindexera le contenu en cas de modification de l'une de ces colonnes
allowedUsersColumns	Entrez le nom de la colonne contenant l'utilisa teur devant IDs être autorisé à accéder au contenu.
allowedGroupsColumn	Entrez le nom de la colonne contenant l'utilisa teur devant IDs être autorisé à accéder au contenu.
source URIColumn	Entrez le nom de la colonne contenant la source URLs à indexer.

Configuration	Description
isSslEnabled	Entrez des instructions de requête SQL telles que les opérations SELECT et JOIN. La taille des requêtes SQL doit être inférieure à 32 Ko. Amazon Kendra analysera tout le contenu de la base de données correspondant à votre requête.
type	Type de source de données. Spécifiez `JDBC` comme type de source de données.

Configuration	Description
Mode de synchronisation	Spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Vous pouvez choisir entre : • `FORCED_FULL_CRAWL` pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index. • `FULL_CRAWL` pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modificat ions de contenu et indexer le contenu modifié depuis la dernière synchronisation. • `CHANGE_LOG` pour indexer uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchroni se avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

Configuration	Description
Secrétaire Arn	Le nom de ressource Amazon (ARN) d'un secret Secrets Manager contenant le nom d'utilisateur et le mot de passe requis pour se connecter à votre base de données. Le secret doit contenir une structure JSON avec les clés suivantes : `{ "user name": "database user name", "password": " password" }`
version	Version du modèle actuellement prise en charge.

Aurora Schéma JSON (PostgreSQL)

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "dbType": {
              "type": "string",
              "enum": [
                "mysql",
                "db2",
                "postgresql",
                "oracle",
                "sqlserver"
              ]
            },
            "dbHost": {
```

```
            "type": "string"
          },
          "dbPort": {
            "type": "string"
          },
          "dbInstance": {
            "type": "string"
          }
        },
        "required": [
          "dbType",
          "dbHost",
          "dbPort",
          "dbInstance"
        ]
      }
    },
    "required": [
      "repositoryEndpointMetadata"
    ]
  },
  "repositoryConfigurations": {
    "type": "object",
    "properties": {
      "document": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string"
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  }
                },
                "required": [
```

```
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      }
    },
    "required": [
    ]
  },
  "additionalProperties": {
    "type": "object",
    "properties": {
      "primaryKey": {
        "type": "string"
      },
      "titleColumn": {
        "type": "string"
      },
      "bodyColumn": {
        "type": "string"
      },
      "sqlQuery": {
        "type": "string",
        "not": {
          "pattern": ";+"
        }
      },
      "timestampColumn": {
        "type": "string"
      },
      "timestampFormat": {
        "type": "string"
      },
      "timezone": {
        "type": "string"
      },
      "changeDetectingColumns": {
```

```
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "allowedUsersColumn": {
          "type": "string"
        },
        "allowedGroupsColumn": {
          "type": "string"
        },
        "sourceURIColumn": {
          "type": "string"
        },
        "isSslEnabled": {
          "type": "boolean"
        }
      },
      "required": ["primaryKey", "titleColumn", "bodyColumn", "sqlQuery"]
    },
    "type" : {
      "type" : "string",
      "pattern": "JDBC"
    },
    "syncMode": {
      "type": "string",
      "enum": [
        "FORCED_FULL_CRAWL",
        "FULL_CRAWL",
        "CHANGE_LOG"
      ]
    },
    "secretArn": {
      "type": "string"
    }
  },
  "version": {
    "type": "string",
    "anyOf": [
      {
        "pattern": "1.0.0"
      }
    ]
  },
```

```
  "required": [
      "connectionConfiguration",
      "repositoryConfigurations",
      "syncMode",
      "additionalProperties",
      "secretArn",
      "type"
  ]
}
```

Amazon RDS Schéma de modèle (Microsoft SQL Server)

Vous incluez un JSON qui contient le schéma de la source de données dans le cadre de l'TemplateConfigurationobjet. Spécifiez le type de source de donnéesJDBC, le type de base de donnéessqlserver, un secret pour vos informations d'authentification et les autres configurations nécessaires. Vous spécifiez ensuite TEMPLATE le Type moment de l'appel CreateDataSource.

Vous pouvez utiliser le modèle fourni dans ce guide du développeur. Consultez Amazon RDS Schéma JSON (Microsoft SQL Server).

Le tableau suivant décrit les paramètres du schéma JSON Amazon RDS (Microsoft SQL Server).

Configuration	Description
Configuration de connexion	Informations de configuration pour le point de terminaison de la source de données.
repositoryEndpointMetadata	Informations de configuration requises pour connecter votre source de données. • DBType : le type de base de données Java que vous utilisez, que ce soit, mysqldb2, postgresql ou. oracle sqlserver • DBHost : nom d'hôte de la base de données. • DBPort : port de base de données. • DBInstance : instance de base de données.
Configurations du référentiel	Informations de configuration pour le contenu de la source de données. Par exemple,

Configuration	Description
	configurer des types spécifiques de contenu et des mappages de champs. Spécifiez le type de source de données et l'ARN secret.
document	Liste d'objets qui mappent les attributs ou les noms de champs du contenu de votre base de données aux noms de champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.
Propriétés supplémentaires	Options de configuration supplémentaires pour votre contenu dans votre source de données. À utiliser pour inclure ou exclure un contenu spécifique dans la source de données de votre base de données.
Clé primaire	Fournissez la clé primaire pour la table de base de données. Cela identifie une table au sein de votre base de données.
Colonne de titre	Indiquez le nom de la colonne du titre du document dans votre table de base de données.
Colonne Body	Indiquez le nom de la colonne du titre du document dans votre table de base de données.
sqlQuery	Entrez des instructions de requête SQL telles que les opérations SELECT et JOIN. La taille des requêtes SQL doit être inférieure à 32 Ko. Amazon Kendra analysera tout le contenu de la base de données correspondant à votre requête.

Configuration	Description
Colonne d'horodatage	Entrez le nom de la colonne contenant les horodatages. Amazon Kendra utilise les informations d'horodatage pour détecter les modifications apportées à votre contenu et synchroniser uniquement le contenu modifié.
Format d'horodatage	Entrez le nom de la colonne contenant les formats d'horodatage à utiliser pour détecter les modifications de contenu et resynchroniser votre contenu.
timezone	Entrez le nom de la colonne contenant les fuseaux horaires du contenu à analyser.
changeDetectingColumns	Entrez le nom des colonnes qui Amazon Kendra seront utilisées pour détecter les modifications de contenu. Amazon Kendra réindexera le contenu en cas de modification de l'une de ces colonnes
allowedUsersColumns	Entrez le nom de la colonne contenant l'utilisa teur devant IDs être autorisé à accéder au contenu.
allowedGroupsColumn	Entrez le nom de la colonne contenant l'utilisa teur devant IDs être autorisé à accéder au contenu.
source URIColumn	Entrez le nom de la colonne contenant la source URLs à indexer.

Configuration	Description
isSslEnabled	Entrez des instructions de requête SQL telles que les opérations SELECT et JOIN. La taille des requêtes SQL doit être inférieure à 32 Ko. Amazon Kendra analysera tout le contenu de la base de données correspondant à votre requête.
type	Type de source de données. Spécifiez `JDBC` comme type de source de données.

Configuration	Description
Mode de synchronisation	Spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Vous pouvez choisir entre : • `FORCED_FULL_CRAWL` pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index. • `FULL_CRAWL` pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modificat ions de contenu et indexer le contenu modifié depuis la dernière synchronisation. • `CHANGE_LOG` pour indexer uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchroni se avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

Configuration	Description
Secrétaire Arn	Le nom de ressource Amazon (ARN) d'un secret Secrets Manager contenant le nom d'utilisateur et le mot de passe requis pour se connecter à votre base de données. Le secret doit contenir une structure JSON avec les clés suivantes : `{ "user name": "database user name", "password": " password" }`
version	Version du modèle actuellement prise en charge.

Amazon RDS Schéma JSON (Microsoft SQL Server)

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "dbType": {
              "type": "string",
              "enum": [
                "mysql",
                "db2",
                "postgresql",
                "oracle",
                "sqlserver"
              ]
            },
            "dbHost": {
```

```
            "type": "string"
          },
          "dbPort": {
            "type": "string"
          },
          "dbInstance": {
            "type": "string"
          }
        },
        "required": [
          "dbType",
          "dbHost",
          "dbPort",
          "dbInstance"
        ]
      }
    },
    "required": [
      "repositoryEndpointMetadata"
    ]
  },
  "repositoryConfigurations": {
    "type": "object",
    "properties": {
      "document": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string"
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  }
                },
                "required": [
```

```
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      }
    },
    "required": [
    ]
  },
  "additionalProperties": {
    "type": "object",
    "properties": {
      "primaryKey": {
        "type": "string"
      },
      "titleColumn": {
        "type": "string"
      },
      "bodyColumn": {
        "type": "string"
      },
      "sqlQuery": {
        "type": "string",
        "not": {
          "pattern": ";+"
        }
      },
      "timestampColumn": {
        "type": "string"
      },
      "timestampFormat": {
        "type": "string"
      },
      "timezone": {
        "type": "string"
      },
      "changeDetectingColumns": {
```

```
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "allowedUsersColumn": {
          "type": "string"
        },
        "allowedGroupsColumn": {
          "type": "string"
        },
        "sourceURIColumn": {
          "type": "string"
        },
        "isSslEnabled": {
          "type": "boolean"
        }
      },
      "required": ["primaryKey", "titleColumn", "bodyColumn", "sqlQuery"]
    },
    "type" : {
      "type" : "string",
      "pattern": "JDBC"
    },
    "syncMode": {
      "type": "string",
      "enum": [
        "FORCED_FULL_CRAWL",
        "FULL_CRAWL",
        "CHANGE_LOG"
      ]
    },
    "secretArn": {
      "type": "string"
    }
  },
  "version": {
    "type": "string",
    "anyOf": [
      {
        "pattern": "1.0.0"
      }
    ]
  },
```

```
  "required": [
      "connectionConfiguration",
      "repositoryConfigurations",
      "syncMode",
      "additionalProperties",
      "secretArn",
      "type"
  ]
}
```

Amazon RDS Schéma de modèle (MySQL)

Vous incluez un JSON qui contient le schéma de la source de données dans le cadre de l'TemplateConfigurationobjet. Spécifiez le type de source de donnéesJDBC, le type de base de donnéesmysql, un secret pour vos informations d'authentification et les autres configurations nécessaires. Vous spécifiez ensuite TEMPLATE le Type moment de l'appel CreateDataSource.

Vous pouvez utiliser le modèle fourni dans ce guide du développeur. Consultez Amazon RDS Schéma JSON (MySQL).

Le tableau suivant décrit les paramètres du schéma JSON Amazon RDS (MySQL).

Configuration	Description
Configuration de connexion	Informations de configuration pour le point de terminaison de la source de données.
repositoryEndpointMetadata	Informations de configuration requises pour connecter votre source de données. • DBType : le type de base de données Java que vous utilisez, que ce soit, mysqldb2, postgresql ou. oracle sqlserver • DBHost : nom d'hôte de la base de données. • DBPort : port de base de données. • DBInstance : instance de base de données.
Configurations du référentiel	Informations de configuration pour le contenu de la source de données. Par exemple,

Configuration	Description
	configurer des types spécifiques de contenu et des mappages de champs. Spécifiez le type de source de données et l'ARN secret.
document	Liste d'objets qui mappent les attributs ou les noms de champs du contenu de votre base de données aux noms de champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.
Propriétés supplémentaires	Options de configuration supplémentaires pour votre contenu dans votre source de données. À utiliser pour inclure ou exclure un contenu spécifique dans la source de données de votre base de données.
Clé primaire	Fournissez la clé primaire pour la table de base de données. Cela identifie une table au sein de votre base de données.
Colonne de titre	Indiquez le nom de la colonne du titre du document dans votre table de base de données.
Colonne Body	Indiquez le nom de la colonne du titre du document dans votre table de base de données.
sqlQuery	Entrez des instructions de requête SQL telles que les opérations SELECT et JOIN. La taille des requêtes SQL doit être inférieure à 32 Ko. Amazon Kendra analysera tout le contenu de la base de données correspondant à votre requête.

Configuration	Description
Colonne d'horodatage	Entrez le nom de la colonne contenant les horodatages. Amazon Kendra utilise les informations d'horodatage pour détecter les modifications apportées à votre contenu et synchroniser uniquement le contenu modifié.
Format d'horodatage	Entrez le nom de la colonne contenant les formats d'horodatage à utiliser pour détecter les modifications de contenu et resynchroniser votre contenu.
timezone	Entrez le nom de la colonne contenant les fuseaux horaires du contenu à analyser.
changeDetectingColumns	Entrez le nom des colonnes qui Amazon Kendra seront utilisées pour détecter les modifications de contenu. Amazon Kendra réindexera le contenu en cas de modification de l'une de ces colonnes
allowedUsersColumns	Entrez le nom de la colonne contenant l'utilisa teur devant IDs être autorisé à accéder au contenu.
allowedGroupsColumn	Entrez le nom de la colonne contenant l'utilisa teur devant IDs être autorisé à accéder au contenu.
source URIColumn	Entrez le nom de la colonne contenant la source URLs à indexer.

Configuration	Description
isSslEnabled	Entrez des instructions de requête SQL telles que les opérations SELECT et JOIN. La taille des requêtes SQL doit être inférieure à 32 Ko. Amazon Kendra analysera tout le contenu de la base de données correspondant à votre requête.
type	Type de source de données. Spécifiez `JDBC` comme type de source de données.

Configuration	Description
Mode de synchronisation	Spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Vous pouvez choisir entre : • `FORCED_FULL_CRAWL` pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index. • `FULL_CRAWL` pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modificat ions de contenu et indexer le contenu modifié depuis la dernière synchronisation. • `CHANGE_LOG` pour indexer uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchroni se avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

Configuration	Description
Secrétaire Arn	Le nom de ressource Amazon (ARN) d'un secret Secrets Manager contenant le nom d'utilisateur et le mot de passe requis pour se connecter à votre base de données. Le secret doit contenir une structure JSON avec les clés suivantes : `{ "user name": "database user name", "password": " password" }`
version	Version du modèle actuellement prise en charge.

Amazon RDS Schéma JSON (MySQL)

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "dbType": {
              "type": "string",
              "enum": [
                "mysql",
                "db2",
                "postgresql",
                "oracle",
                "sqlserver"
              ]
            },
            "dbHost": {
```

```
              "type": "string"
            },
            "dbPort": {
              "type": "string"
            },
            "dbInstance": {
              "type": "string"
            }
          },
          "required": [
            "dbType",
            "dbHost",
            "dbPort",
            "dbInstance"
          ]
        }
      },
      "required": [
        "repositoryEndpointMetadata"
      ]
    },
    "repositoryConfigurations": {
      "type": "object",
      "properties": {
        "document": {
          "type": "object",
          "properties": {
            "fieldMappings": {
              "type": "array",
              "items": [
                {
                  "type": "object",
                  "properties": {
                    "indexFieldName": {
                      "type": "string"
                    },
                    "indexFieldType": {
                      "type": "string"
                    },
                    "dataSourceFieldName": {
                      "type": "string"
                    }
                  },
                  "required": [
```

```
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      }
    },
    "required": [
    ]
  },
  "additionalProperties": {
    "type": "object",
    "properties": {
      "primaryKey": {
        "type": "string"
      },
      "titleColumn": {
        "type": "string"
      },
      "bodyColumn": {
        "type": "string"
      },
      "sqlQuery": {
        "type": "string",
        "not": {
          "pattern": ";+"
        }
      },
      "timestampColumn": {
        "type": "string"
      },
      "timestampFormat": {
        "type": "string"
      },
      "timezone": {
        "type": "string"
      },
      "changeDetectingColumns": {
```

```
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "allowedUsersColumn": {
          "type": "string"
        },
        "allowedGroupsColumn": {
          "type": "string"
        },
        "sourceURIColumn": {
          "type": "string"
        },
        "isSslEnabled": {
          "type": "boolean"
        }
      },
      "required": ["primaryKey", "titleColumn", "bodyColumn", "sqlQuery"]
    },
    "type" : {
      "type" : "string",
      "pattern": "JDBC"
    },
    "syncMode": {
      "type": "string",
      "enum": [
        "FORCED_FULL_CRAWL",
        "FULL_CRAWL",
        "CHANGE_LOG"
      ]
    },
    "secretArn": {
      "type": "string"
    }
  },
  "version": {
    "type": "string",
    "anyOf": [
      {
        "pattern": "1.0.0"
      }
    ]
  },
```

```
  "required": [
      "connectionConfiguration",
      "repositoryConfigurations",
      "syncMode",
      "additionalProperties",
      "secretArn",
      "type"
  ]
}
```

Amazon RDS schéma de modèle (Oracle)

Vous incluez un JSON qui contient le schéma de la source de données dans le cadre de l'TemplateConfigurationobjet. Spécifiez le type de source de donnéesJDBC, le type de base de donnéesoracle, un secret pour vos informations d'authentification et les autres configurations nécessaires. Vous spécifiez ensuite TEMPLATE le Type moment de l'appel CreateDataSource.

Vous pouvez utiliser le modèle fourni dans ce guide du développeur. Consultez Amazon RDS Schéma JSON (Oracle).

Le tableau suivant décrit les paramètres du schéma JSON Amazon RDS (Oracle).

Configuration	Description
Configuration de connexion	Informations de configuration pour le point de terminaison de la source de données.
repositoryEndpointMetadata	Informations de configuration requises pour connecter votre source de données. • DBType : le type de base de données Java que vous utilisez, que ce soit, mysqldb2, postgresql ou. oracle sqlserver • DBHost : nom d'hôte de la base de données. • DBPort : port de base de données. • DBInstance : instance de base de données.
Configurations du référentiel	Informations de configuration pour le contenu de la source de données. Par exemple,

Configuration	Description
	configurer des types spécifiques de contenu et des mappages de champs. Spécifiez le type de source de données et l'ARN secret.
document	Liste d'objets qui mappent les attributs ou les noms de champs du contenu de votre base de données aux noms de champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.
Propriétés supplémentaires	Options de configuration supplémentaires pour votre contenu dans votre source de données. À utiliser pour inclure ou exclure un contenu spécifique dans la source de données de votre base de données.
Clé primaire	Fournissez la clé primaire pour la table de base de données. Cela identifie une table au sein de votre base de données.
Colonne de titre	Indiquez le nom de la colonne du titre du document dans votre table de base de données.
Colonne Body	Indiquez le nom de la colonne du titre du document dans votre table de base de données.
sqlQuery	Entrez des instructions de requête SQL telles que les opérations SELECT et JOIN. La taille des requêtes SQL doit être inférieure à 32 Ko. Amazon Kendra analysera tout le contenu de la base de données correspondant à votre requête.

Configuration	Description
Colonne d'horodatage	Entrez le nom de la colonne contenant les horodatages. Amazon Kendra utilise les informations d'horodatage pour détecter les modifications apportées à votre contenu et synchroniser uniquement le contenu modifié.
Format d'horodatage	Entrez le nom de la colonne contenant les formats d'horodatage à utiliser pour détecter les modifications de contenu et resynchroniser votre contenu.
timezone	Entrez le nom de la colonne contenant les fuseaux horaires du contenu à analyser.
changeDetectingColumns	Entrez le nom des colonnes qui Amazon Kendra seront utilisées pour détecter les modifications de contenu. Amazon Kendra réindexera le contenu en cas de modification de l'une de ces colonnes
allowedUsersColumns	Entrez le nom de la colonne contenant l'utilisa teur devant IDs être autorisé à accéder au contenu.
allowedGroupsColumn	Entrez le nom de la colonne contenant l'utilisa teur devant IDs être autorisé à accéder au contenu.
source URIColumn	Entrez le nom de la colonne contenant la source URLs à indexer.

Configuration	Description
isSslEnabled	Entrez des instructions de requête SQL telles que les opérations SELECT et JOIN. La taille des requêtes SQL doit être inférieure à 32 Ko. Amazon Kendra analysera tout le contenu de la base de données correspondant à votre requête.
type	Type de source de données. Spécifiez `JDBC` comme type de source de données.

Configuration	Description
Mode de synchronisation	Spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Vous pouvez choisir entre : • `FORCED_FULL_CRAWL` pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index. • `FULL_CRAWL` pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modificat ions de contenu et indexer le contenu modifié depuis la dernière synchronisation. • `CHANGE_LOG` pour indexer uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchroni se avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

Configuration	Description
Secrétaire Arn	Le nom de ressource Amazon (ARN) d'un secret Secrets Manager contenant le nom d'utilisateur et le mot de passe requis pour se connecter à votre base de données. Le secret doit contenir une structure JSON avec les clés suivantes : `{ "user name": "database user name", "password": " password" }`
version	Version du modèle actuellement prise en charge.

Amazon RDS Schéma JSON (Oracle)

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "dbType": {
              "type": "string",
              "enum": [
                "mysql",
                "db2",
                "postgresql",
                "oracle",
                "sqlserver"
              ]
            },
            "dbHost": {
```

```
            "type": "string"
          },
          "dbPort": {
            "type": "string"
          },
          "dbInstance": {
            "type": "string"
          }
        },
        "required": [
          "dbType",
          "dbHost",
          "dbPort",
          "dbInstance"
        ]
      }
    },
    "required": [
      "repositoryEndpointMetadata"
    ]
  },
  "repositoryConfigurations": {
    "type": "object",
    "properties": {
      "document": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string"
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  }
                },
                "required": [
```

```
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      }
    },
    "required": [
    ]
  },
  "additionalProperties": {
    "type": "object",
    "properties": {
      "primaryKey": {
        "type": "string"
      },
      "titleColumn": {
        "type": "string"
      },
      "bodyColumn": {
        "type": "string"
      },
      "sqlQuery": {
        "type": "string",
        "not": {
          "pattern": ";+"
        }
      },
      "timestampColumn": {
        "type": "string"
      },
      "timestampFormat": {
        "type": "string"
      },
      "timezone": {
        "type": "string"
      },
      "changeDetectingColumns": {
```

```
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "allowedUsersColumn": {
          "type": "string"
        },
        "allowedGroupsColumn": {
          "type": "string"
        },
        "sourceURIColumn": {
          "type": "string"
        },
        "isSslEnabled": {
          "type": "boolean"
        }
      },
      "required": ["primaryKey", "titleColumn", "bodyColumn", "sqlQuery"]
    },
    "type" : {
      "type" : "string",
      "pattern": "JDBC"
    },
    "syncMode": {
      "type": "string",
      "enum": [
        "FORCED_FULL_CRAWL",
        "FULL_CRAWL",
        "CHANGE_LOG"
      ]
    },
    "secretArn": {
      "type": "string"
    }
  },
  "version": {
    "type": "string",
    "anyOf": [
      {
        "pattern": "1.0.0"
      }
    ]
  },
```

```
  "required": [
      "connectionConfiguration",
      "repositoryConfigurations",
      "syncMode",
      "additionalProperties",
      "secretArn",
      "type"
  ]
}
```

Amazon RDS Schéma de modèle (PostgreSQL)

Vous incluez un JSON qui contient le schéma de la source de données dans le cadre de l'TemplateConfigurationobjet. Spécifiez le type de source de donnéesJDBC, le type de base de donnéespostgresql, un secret pour vos informations d'authentification et les autres configurations nécessaires. Vous spécifiez ensuite TEMPLATE le Type moment de l'appel CreateDataSource.

Vous pouvez utiliser le modèle fourni dans ce guide du développeur. Consultez Amazon RDS Schéma JSON (PostgreSQL).

Le tableau suivant décrit les paramètres du schéma Amazon RDS JSON (PostgreSQL).

Configuration	Description
Configuration de connexion	Informations de configuration pour le point de terminaison de la source de données.
repositoryEndpointMetadata	Informations de configuration requises pour connecter votre source de données. • DBType : le type de base de données Java que vous utilisez, que ce soit, mysqldb2, postgresql ou. oracle sqlserver • DBHost : nom d'hôte de la base de données. • DBPort : port de base de données. • DBInstance : instance de base de données.
Configurations du référentiel	Informations de configuration pour le contenu de la source de données. Par exemple,

Configuration	Description
	configurer des types spécifiques de contenu et des mappages de champs. Spécifiez le type de source de données et l'ARN secret.
document	Liste d'objets qui mappent les attributs ou les noms de champs du contenu de votre base de données aux noms de champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.
Propriétés supplémentaires	Options de configuration supplémentaires pour votre contenu dans votre source de données. À utiliser pour inclure ou exclure un contenu spécifique dans la source de données de votre base de données.
Clé primaire	Fournissez la clé primaire pour la table de base de données. Cela identifie une table au sein de votre base de données.
Colonne de titre	Indiquez le nom de la colonne du titre du document dans votre table de base de données.
Colonne Body	Indiquez le nom de la colonne du titre du document dans votre table de base de données.
sqlQuery	Entrez des instructions de requête SQL telles que les opérations SELECT et JOIN. La taille des requêtes SQL doit être inférieure à 32 Ko. Amazon Kendra analysera tout le contenu de la base de données correspondant à votre requête.

Configuration	Description
Colonne d'horodatage	Entrez le nom de la colonne contenant les horodatages. Amazon Kendra utilise les informations d'horodatage pour détecter les modifications apportées à votre contenu et synchroniser uniquement le contenu modifié.
Format d'horodatage	Entrez le nom de la colonne contenant les formats d'horodatage à utiliser pour détecter les modifications de contenu et resynchroniser votre contenu.
timezone	Entrez le nom de la colonne contenant les fuseaux horaires du contenu à analyser.
changeDetectingColumns	Entrez le nom des colonnes qui Amazon Kendra seront utilisées pour détecter les modifications de contenu. Amazon Kendra réindexera le contenu en cas de modification de l'une de ces colonnes
allowedUsersColumns	Entrez le nom de la colonne contenant l'utilisa teur devant IDs être autorisé à accéder au contenu.
allowedGroupsColumn	Entrez le nom de la colonne contenant l'utilisa teur devant IDs être autorisé à accéder au contenu.
source URIColumn	Entrez le nom de la colonne contenant la source URLs à indexer.

Configuration	Description
isSslEnabled	Entrez des instructions de requête SQL telles que les opérations SELECT et JOIN. La taille des requêtes SQL doit être inférieure à 32 Ko. Amazon Kendra analysera tout le contenu de la base de données correspondant à votre requête.
type	Type de source de données. Spécifiez `JDBC` comme type de source de données.

Configuration	Description
Mode de synchronisation	Spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Vous pouvez choisir entre : • `FORCED_FULL_CRAWL` pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index. • `FULL_CRAWL` pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modificat ions de contenu et indexer le contenu modifié depuis la dernière synchronisation. • `CHANGE_LOG` pour indexer uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchroni se avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

Configuration	Description
Secrétaire Arn	Le nom de ressource Amazon (ARN) d'un secret Secrets Manager contenant le nom d'utilisateur et le mot de passe requis pour se connecter à votre base de données. Le secret doit contenir une structure JSON avec les clés suivantes : `{ "user name": "database user name", "password": " password" }`
version	Version du modèle actuellement prise en charge.

Amazon RDS Schéma JSON (PostgreSQL)

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "dbType": {
              "type": "string",
              "enum": [
                "mysql",
                "db2",
                "postgresql",
                "oracle",
                "sqlserver"
              ]
            },
            "dbHost": {
```

```
            "type": "string"
          },
          "dbPort": {
            "type": "string"
          },
          "dbInstance": {
            "type": "string"
          }
        },
        "required": [
          "dbType",
          "dbHost",
          "dbPort",
          "dbInstance"
        ]
      }
    },
    "required": [
      "repositoryEndpointMetadata"
    ]
  },
  "repositoryConfigurations": {
    "type": "object",
    "properties": {
      "document": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string"
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  }
                },
                "required": [
```

```
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      }
    },
    "required": [
    ]
  },
  "additionalProperties": {
    "type": "object",
    "properties": {
      "primaryKey": {
        "type": "string"
      },
      "titleColumn": {
        "type": "string"
      },
      "bodyColumn": {
        "type": "string"
      },
      "sqlQuery": {
        "type": "string",
        "not": {
          "pattern": ";+"
        }
      },
      "timestampColumn": {
        "type": "string"
      },
      "timestampFormat": {
        "type": "string"
      },
      "timezone": {
        "type": "string"
      },
      "changeDetectingColumns": {
```

```
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "allowedUsersColumn": {
          "type": "string"
        },
        "allowedGroupsColumn": {
          "type": "string"
        },
        "sourceURIColumn": {
          "type": "string"
        },
        "isSslEnabled": {
          "type": "boolean"
        }
      },
      "required": ["primaryKey", "titleColumn", "bodyColumn", "sqlQuery"]
    },
    "type" : {
      "type" : "string",
      "pattern": "JDBC"
    },
    "syncMode": {
      "type": "string",
      "enum": [
        "FORCED_FULL_CRAWL",
        "FULL_CRAWL",
        "CHANGE_LOG"
      ]
    },
    "secretArn": {
      "type": "string"
    }
  },
  "version": {
    "type": "string",
    "anyOf": [
      {
        "pattern": "1.0.0"
      }
    ]
  },
```

```
  "required": [
      "connectionConfiguration",
      "repositoryConfigurations",
      "syncMode",
      "additionalProperties",
      "secretArn",
      "type"
  ]
}
```

Amazon S3 schéma de modèle

Vous incluez un JSON contenant le schéma de source de données dans le cadre de la configuration du modèle. Vous fournissez le nom du compartiment S3 dans le cadre de la configuration de la connexion ou des détails du point de terminaison du référentiel. Spécifiez également le type de source de S3 données et les autres configurations nécessaires. Vous spécifiez ensuite TEMPLATE le Type moment de l'appel CreateDataSource.

Vous pouvez utiliser le modèle fourni dans ce guide du développeur. Consultez Schéma JSON S3.

Le tableau suivant décrit les paramètres du schéma Amazon S3 JSON.

Configuration	Description
Configuration de connexion	Informations de configuration pour le point de terminaison de la source de données.
repositoryEndpointMetadata	Informations sur le point de terminaison de la source de données.
BucketName	Le nom de votre Amazon S3 compartiment.
Configurations du référentiel	Informations de configuration pour le contenu de la source de données. Par exemple, configurer des types spécifiques de contenu et des mappages de champs.
Propriétés supplémentaires	Options de configuration supplémentaires pour votre contenu dans votre source de données

Configuration	Description
• Modèles d'inclusion • Schémas d'exclusion • Préfixes d'inclusion • Préfixes d'exclusion	Liste de modèles d'expressions régulières permettant d'inclure ou d'exclure des fichiers spécifiques dans votre source de Amazon S3 données. Les fichiers qui correspondent aux modèles sont inclus dans l'index. Les fichiers qui ne correspondent pas aux modèles sont exclus de l'index. Si un fichier correspond à la fois à un modèle d'inclusion et d'exclusion, le modèle d'exclusion a la priorité et le fichier n'est pas inclus dans l'index.
aclConfigurationFileChemin	Le chemin du fichier qui contrôle l'accès aux documents d'un Amazon Kendra index.
metadataFilesPrefix	Emplacement des fichiers de métadonnées au sein de votre compartiment.
Mode de synchronisation	Spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Vous pouvez choisir entre : • `FORCED_FULL_CRAWL` pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index. • `FULL_CRAWL` pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modificat ions de contenu et indexer le contenu modifié depuis la dernière synchronisation.

Configuration	Description
type	Type de source de données. Spécifiez S3 comme type de source de données.
version	Version du modèle prise en charge.

Schéma JSON S3

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "BucketName": {
              "type": "string"
            }
          },
          "required": [
            "BucketName"
          ]
        }
      },
      "required": [
        "repositoryEndpointMetadata"
      ]
    },
    "repositoryConfigurations": {
      "type": "object",
      "properties": {
        "document": {
          "type": "object",
          "properties": {
            "fieldMappings": {
              "type": "array",
              "items": [
                {
```

```
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string",
                    "enum": [
                      "STRING"
                    ]
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  }
                },
                "required": [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      }
    },
    "required": [
      "document"
    ]
  },
  "additionalProperties": {
    "type": "object",
    "properties": {
      "inclusionPatterns": {
        "type": "array"
      },
      "exclusionPatterns": {
        "type": "array"
      },
      "inclusionPrefixes": {
        "type": "array"
```

```
        },
        "exclusionPrefixes": {
          "type": "array"
        },
        "aclConfigurationFilePath": {
          "type": "string"
        },
        "metadataFilesPrefix": {
          "type": "string"
        }
      }
    },
    "syncMode": {
      "type": "string",
      "enum": [
        "FULL_CRAWL",
        "FORCED_FULL_CRAWL"
      ]
    },
    "type": {
      "type": "string",
      "pattern": "S3"
    },
    "version": {
      "type": "string",
      "anyOf": [
        {
          "pattern": "1.0.0"
        }
      ]
    }
  },
  "required": [
    "connectionConfiguration",
    "type",
    "syncMode",
    "repositoryConfigurations"
  ]
}
```

Amazon Kendra Schéma du modèle Web Crawler

Vous incluez un JSON qui contient le schéma de la source de données dans le cadre de l'TemplateConfigurationobjet.

Vous fournissez le point de départ ou le point de départ URLs, ou vous pouvez fournir le plan du site URLs, dans le cadre de la configuration de la connexion ou des détails du point de terminaison du référentiel. Au lieu de répertorier manuellement tous vos fichiers URLs, vous pouvez indiquer le chemin d'accès au Amazon S3 compartiment qui stocke un fichier texte pour votre liste de fichiers XML de départ URLs ou de plan de site, que vous pouvez regrouper dans un fichier ZIP dans S3.

Vous spécifiez également le type de source de donnéesWEBCRAWLERV2, les informations d'authentification du site Web et le type d'authentification si vos sites Web nécessitent une authentification, ainsi que les autres configurations nécessaires.

Vous spécifiez ensuite TEMPLATE le Type moment de l'appel CreateDataSource.

> **Important**
>
> La création de connecteurs Web Crawler v2.0 n'est pas prise en charge par. AWS CloudFormation Utilisez le connecteur Web Crawler v1.0 si vous avez besoin d' AWS CloudFormation assistance.

Lorsque vous sélectionnez des sites web à indexer, vous devez respecter les Politiques d'Amazon en matière d'utilisation acceptable et toutes les autres conditions d'Amazon. N'oubliez pas que vous ne devez utiliser Amazon Kendra Web Crawler que pour indexer vos propres pages Web ou les pages Web que vous êtes autorisé à indexer. Pour savoir comment empêcher Amazon Kendra Web Crawler d'indexer vos sites Web, consultez. Configuration du *robots.txt* fichier pour Amazon Kendra Web Crawler

Vous pouvez utiliser le modèle fourni dans ce guide du développeur. Consultez Amazon Kendra Schéma JSON de Web Crawler.

Le tableau suivant décrit les paramètres du schéma JSON du Amazon Kendra Web Crawler.

Configuration	Description
Configuration de connexion	Informations de configuration pour le point de terminaison de la source de données.
repositoryEndpointMetadata	Informations sur le point de terminaison de la source de données.
siteMapUrls	La liste URLs des plans du site Web que vous souhaitez explorer. Vous pouvez répertorier jusqu'à trois plans de site URLs.
s3 SeedUrl	Le chemin S3 vers le fichier texte qui stocke la liste des points de départ ou de départ URLs. Par exemple, s3://bucket-name/directory/. Chaque URL du fichier texte doit être formatée sur une ligne distincte. Vous pouvez répertorier jusqu'à 100 graines URLs dans un fichier.
s3 SiteMapUrl	Le chemin S3 vers les fichiers XML du plan du site. Par exemple, s3://bucket-name/directory/ . Vous pouvez répertorier jusqu'à trois fichiers XML de plan de site. Vous pouvez regrouper plusieurs fichiers de plan du site dans un fichier ZIP et le stocker dans votre Amazon S3 compartiment.
seedUrlConnections	La liste des graines ou du point de départ URLs des sites Web que vous souhaitez explorer. Vous pouvez répertorier jusqu'à 100 graines. URLs
Voir l'URL	URL du point de départ ou de départ.
authentification	Type d'authentification si vos sites Web nécessitent la même authentification, sinon spécifiez-leNoAuthentication .

Configuration	Description
Configurations du référentiel	Informations de configuration pour le contenu de la source de données. Par exemple, configurer des types spécifiques de contenu et des mappages de champs.
• page Web • attachment	Liste d'objets qui mappent les attributs ou les noms de champs de vos pages Web et de vos fichiers de pages Web aux noms de champs d'Amazon Kendra index. Par exemple, la balise de titre de page Web HTML peut être mappée au champ d'`_document_title` index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.
Mode de synchronisation	Spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Vous pouvez choisir entre : • `FORCED_FULL_CRAWL` pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index. • `FULL_CRAWL` pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

Configuration	Description
Propriétés supplémentaires	Options de configuration supplémentaires pour votre contenu dans votre source de données.
Limite de taux	Le nombre maximum d' URLs explorations par hôte de site Web par minute.
maxFileSize	Taille maximale (en Mo) d'une page Web ou d'une pièce jointe à analyser.
Profondeur du crawl	Le nombre de niveaux à partir de l'URL de départ à explorer. Par exemple, la page URL initiale est de profondeur 1 et tous les hyperlien s de cette page qui sont également explorés ont une profondeur de 2.
maxLinksPerURL	Le nombre maximum de pages Web URLs à inclure lors de l'exploration d'un site Web. Ce nombre est indiqué par page Web. Lorsque les pages Web d'un site Web sont explorées, toutes celles vers URLs lesquelles les pages Web renvoient sont également explorées. URLs sur une page Web sont explorés par ordre d'apparition.
crawlSubDomain	`true`pour explorer les domaines du site Web à l'aide de sous-domaines. Par exemple, si l'URL initiale est abc.example.com « », alors « a.abc.example.com » et « b.abc.exa mple.com » sont également analysés. Si vous ne le configurez pas `crawlSubDomain` ou `crawlAllDomain` ne le faites pas`true`, il explore Amazon Kendra uniquement les domaines des sites Web que vous souhaitez explorer.

Configuration	Description
crawlAllDomain	`true`pour explorer les domaines du site Web avec des sous-domaines et d'autres domaines vers lesquels les pages Web renvoient. Si vous ne le configurez pas `crawlSubDomain` ou `crawlAllDomain` ne le faites pas`true`, il explore Amazon Kendra uniquement les domaines des sites Web que vous souhaitez explorer.
Honor Robots	`true`pour respecter les directives robots.txt des sites Web que vous souhaitez explorer. Ces directives contrôlent la manière dont Amazon Kendra Web Crawler explore les sites Web, qu'il soit en Amazon Kendra mesure d'explore r uniquement du contenu spécifique ou de ne pas en explorer un.
Accessoires Crawl	`true`pour explorer les fichiers vers lesquels les pages Web renvoient.
• URLCrawlModèles d'inclusion • URLIndexModèles d'inclusion	Une liste de modèles d'expressions régulière s incluant l'exploration de certains hyperliens URLs et l'indexation de tous les hyperliens sur ces pages Web URL. URLs qui correspondent aux modèles sont inclus dans l'index. URLs qui ne correspondent pas aux modèles sont exclus de l'index. Si une URL correspond à la fois à un modèle d'inclusion et à un modèle d'exclusi on, le modèle d'exclusion est prioritaire et les pages Web de l'URL ou du site Web ne sont pas incluses dans l'index.

Configuration	Description
• URLCrawlModèles d'exclusion • URLIndexModèles d'exclusion	Une liste de modèles d'expressions régulière s pour exclure l'exploration de certains liens hypertexte URLs et l'indexation de tout hyperlien sur ces pages Web URL. URLs qui correspondent aux modèles sont exclus de l'index. URLs qui ne correspondent pas aux modèles sont inclus dans l'index. Si une URL correspond à la fois à un modèle d'inclusion et à un modèle d'exclusion, le modèle d'exclusion est prioritaire et les pages Web de l'URL ou du site Web ne sont pas incluses dans l'index.
inclusionFileIndexMotifs	Liste de modèles d'expressions régulières pour inclure certains fichiers de pages Web. Les fichiers qui correspondent aux modèles sont inclus dans l'index. Les fichiers qui ne correspondent pas aux modèles sont exclus de l'index. Si un fichier correspond à la fois à un modèle d'inclusion et à un modèle d'exclusion, le modèle d'exclusion est prioritaire et le fichier n'est pas inclus dans l'index.
exclusionFileIndexMotifs	Liste de modèles d'expressions régulière s permettant d'exclure certains fichiers de pages Web. Les fichiers qui correspondent aux modèles sont exclus de l'index. Les fichiers qui ne correspondent pas aux modèles sont inclus dans l'index. Si un fichier correspond à la fois à un modèle d'inclusion et à un modèle d'exclusi on, le modèle d'exclusion est prioritaire et le fichier n'est pas inclus dans l'index.

Configuration	Description
implicitWaitDuration	implicitWaitDuration indique le temps d'attente du connecteur, en secondes, avant d'explorer une page Web. Fourchette : 0-10 par exemple. « : implicitWaitDuration « 5 »
proxy	Informations de configuration requises pour se connecter à vos sites web internes via un proxy web.
hôte	Le nom d'hôte du serveur proxy que vous souhaitez utiliser pour vous connecter aux sites Web internes. Par exemple, le nom d'hôte de https://a.example.com/page1.htmlest « a.example.com ».
port	Numéro de port du serveur proxy que vous souhaitez utiliser pour vous connecter aux sites Web internes. Par exemple, 443 est le port standard pour HTTPS.
SecretArn (proxy)	Si des informations d'identification de proxy Web sont requises pour se connecter à un hébergeur de site Web, vous pouvez créer un AWS Secrets Manager secret qui stocke les informations d'identification. Indiquez le nom de ressource Amazon (ARN) du secret.
type	Type de source de données. Spécifiez `WEBCRAWLERV2` comme type de source de données.

Configuration	Description
Secrétaire Arn	Le nom de ressource Amazon (ARN) d'un AWS Secrets Manager secret utilisé si vos sites Web nécessitent une authentification pour y accéder. Vous stockez les informations d'authentification du site Web dans le secret qui contient les paires clé-valeur JSON. Si vous utilisez Basic ou NTML/Kerberos, entrez le nom d'utilisateur et le mot de passe. Les clés JSON du secret doivent être `userName` et`password`. Le protocole d'authentification NTLM inclut le hachage des mots de passe et le protocole d'authentification Kerberos inclut le chiffrement des mots de passe. Si vous utilisez l'authentification SAML ou par formulaire, entrez le nom d'utilisateur et le mot de passe, XPath pour le champ du nom d'utilisateur (et le bouton du nom d'utilisateur si vous utilisez le protocole SAML), XPaths pour le champ et le bouton du mot de passe, ainsi que l'URL de la page de connexion . Les clés JSON contenues dans le secret doivent être `userName` `passworduserNameFieldXpath` ,`userNameButtonXpath` ,`passwordFieldXpath` ,`passwordButtonXpath` , et`loginPageUrl` . Vous pouvez trouver le XPaths (langage de chemin XML) des éléments à l'aide des outils de développement de votre navigateur Web. XPaths suivent généralement ce format :`//tagname[@Attribute='Value']` .

Configuration	Description
	Amazon Kendra vérifie également si les informations du point de terminaison (graine URLs) incluses dans le secret sont les mêmes que celles spécifiées dans les détails de configuration du point de terminaison de votre source de données.
version	Version de ce modèle actuellement prise en charge.

Amazon Kendra Schéma JSON de Web Crawler

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "siteMapUrls": {
              "type": "array",
              "items":{
                "type": "string",
                "pattern": "https://.*"
              }
            },
            "s3SeedUrl": {
              "type": "string",
              "pattern": "s3:.*"
            },
            "s3SiteMapUrl": {
              "type": "string",
              "pattern": "s3:.*"
            },
            "seedUrlConnections": {
              "type": "array",
```

```
            "items": [
              {
                "type": "object",
                "properties": {
                  "seedUrl":{
                    "type": "string",
                    "pattern": "https://.*"
                  }
                },
                "required": [
                  "seedUrl"
                ]
              }
            ]
          },
          "authentication": {
            "type": "string",
            "enum": [
              "NoAuthentication",
              "BasicAuth",
              "NTLM_Kerberos",
              "Form",
              "SAML"
            ]
          }
        }
      }
    },
    "required": [
      "repositoryEndpointMetadata"
    ]
  },
  "repositoryConfigurations": {
    "type": "object",
    "properties": {
      "webPage": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
```

```
                "indexFieldName": {
                  "type": "string"
                },
                "indexFieldType": {
                  "type": "string",
                  "enum": [
                    "STRING",
                    "DATE",
                    "LONG"
                  ]
                },
                "dataSourceFieldName": {
                  "type": "string"
                },
                "dateFieldFormat": {
                  "type": "string",
                  "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                }
              },
              "required": [
                "indexFieldName",
                "indexFieldType",
                "dataSourceFieldName"
              ]
            }
          ]
        }
      },
      "required": [
        "fieldMappings"
      ]
    },
    "attachment": {
      "type": "object",
      "properties": {
        "fieldMappings": {
          "type": "array",
          "items": [
            {
              "type": "object",
              "properties": {
                "indexFieldName": {
                  "type": "string"
                },
```

```
                  "indexFieldType": {
                    "type": "string",
                    "enum": [
                      "STRING",
                      "DATE",
                      "LONG"
                    ]
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  },
                  "dateFieldFormat": {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required": [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      }
    }
  },
  "syncMode": {
    "type": "string",
    "enum": [
      "FORCED_FULL_CRAWL",
      "FULL_CRAWL"
    ]
  },
  "additionalProperties": {
    "type": "object",
    "properties": {
      "rateLimit": {
        "type": "string",
        "default": "300"
```

```
        },
        "maxFileSize": {
          "type": "string",
          "default": "50"
        },
        "crawlDepth": {
          "type": "string",
          "default": "2"
        },
        "maxLinksPerUrl": {
          "type": "string",
          "default": "100"
        },
        "crawlSubDomain": {
          "type": "boolean",
          "default": false
        },
        "crawlAllDomain": {
          "type": "boolean",
          "default": false
        },
        "honorRobots": {
          "type": "boolean",
          "default": false
        },
        "crawlAttachments": {
          "type": "boolean",
          "default": false
        },
        "inclusionURLCrawlPatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "exclusionURLCrawlPatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "inclusionURLIndexPatterns": {
          "type": "array",
          "items": {
```

```
            "type": "string"
          }
        },
        "exclusionURLIndexPatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "inclusionFileIndexPatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "exclusionFileIndexPatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "proxy": {
          "type": "object",
          "properties": {
            "host": {
              "type": "string"
            },
            "port": {
              "type": "string"
            },
            "secretArn": {
              "type": "string",
              "minLength": 20,
              "maxLength": 2048
            }
          }
        }
      },
      "implicitWaitDuration":  {
          "type":"object",
          "properties": {
            "innerNumber" : {
              "type": "number",
              "minimum": 0,
```

```
              "maximum": 10
            }
          }
        },
      "required": [
        "rateLimit",
        "maxFileSize",
        "crawlDepth",
        "crawlSubDomain",
        "crawlAllDomain",
        "maxLinksPerUrl",
        "honorRobots"
      ]
    },
    "type": {
      "type": "string",
      "pattern": "WEBCRAWLERV2"
    },
    "secretArn": {
      "type": "string",
      "minLength": 20,
      "maxLength": 2048
    }
  },
  "version": {
    "type": "string",
    "anyOf": [
      {
        "pattern": "1.0.0"
      }
    ]
  },
  "required": [
    "connectionConfiguration",
    "repositoryConfigurations",
    "syncMode",
    "type",
    "additionalProperties"
  ]
}
```

Schéma du modèle Confluence

Vous incluez un JSON qui contient le schéma de la source de données dans le cadre de l'TemplateConfigurationobjet. Vous fournissez l'URL de l'hôte Confluence, la méthode d'hébergement et le type d'authentification dans le cadre de la configuration de la connexion ou des détails du point de terminaison du référentiel. Spécifiez également le type de source de donnéesCONFLUENCEV2, un secret pour vos informations d'authentification et les autres configurations nécessaires. Vous spécifiez ensuite TEMPLATE le Type moment de l'appel CreateDataSource.

Vous pouvez utiliser le modèle fourni dans ce guide du développeur. Consultez Schéma JSON Confluence.

Le tableau suivant décrit les paramètres du schéma Confluence JSON.

Configuration	Description
Configuration de connexion	Informations de configuration pour le point de terminaison de la source de données.
repositoryEndpointMetadata	Informations sur le point de terminaison de la source de données.
URL de l'hôte	URL de votre instance Confluence. Par exemple, *https://example.confluence.com* .
type	La méthode d'hébergement de votre instance Confluence, si SAAS etON_PREM.
authType	La méthode d'authentification de votre instance Confluence, que ce soit BasicOAuth2, ouPersonal-token .
Configurations du référentiel	Informations de configuration pour le contenu de la source de données. Par exemple, configurer des types spécifiques de contenu et des mappages de champs.
• espace	Liste d'objets qui mappent les attributs ou les noms de champs de vos espaces,

Configuration	Description
• page • bloguer • comment • attachment	pages, blogs, commentaires et pièces jointes Confluence pour Amazon Kendra indexer les noms de champs. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données. Les noms des champs de source de données Confluenc e doivent figurer dans vos métadonnées personnalisées Confluence.
Propriétés supplémentaires	Options de configuration supplémentaires pour votre contenu dans votre source de données.
isCrawlAcl	Configurez `true` pour analyser les informati ons de la liste de contrôle d'accès (ACL) de vos documents, si vous disposez d'une ACL et que vous souhaitez l'utiliser pour le contrôle d'accès. Notez que l'ACL spécifie les documents auxquels les utilisateurs et les groupes peuvent accéder. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisa teur ou de son groupe aux documents. Cela signifie que si cette `isCrawlACL` option est désactivée, les documents peuvent être consultés publiquement. Pour plus d'informa tions, consultez la section Filtrage du contexte utilisateur.
fieldForUserID	Spécifiez `email` si vous souhaitez utiliser l'adresse e-mail de l'utilisateur comme nom d'utilisateur. `email`est utilisé par défaut et est actuellement le seul type d'ID utilisateur pris en charge.

Configuration	Description
• inclusionSpaceKeyFiltre • exclusionSpaceKeyFiltre • pageTitleRegEX • blogTitleRegEX • commentTitleRegEX • attachmentTitleRegEX • inclusionFileTypeMotifs • exclusionFileTypeMotifs • inclusionUrlPatterns • exclusionUrlPatterns	Liste de modèles d'expressions régulières à inclure et à and/or exclure certains fichiers de votre source de données Confluence. Les fichiers qui correspondent aux modèles sont inclus dans l'index. Les fichiers qui ne correspondent pas aux modèles sont exclus de l'index. Si un fichier correspond à la fois à un modèle d'inclusion et d'exclusion, le modèle d'exclusion a la priorité et le fichier n'est pas inclus dans l'index.
Hôte proxy	Le nom d'hôte du proxy Web que vous utilisez, sans le `https://` protocole `http://` or.
Port proxy	Numéro de port utilisé par le protocole de transport d'URL de l'hôte. Il doit s'agir d'une valeur numérique comprise entre 0 et 65535.
• isCrawlPersonalEspace • isCrawlArchivedEspace • isCrawlArchivedPage • isCrawlPage • isCrawlBlog • isCrawlPageCommentaire • isCrawlPagePièce jointe • isCrawlBlogCommentaire • isCrawlBlogPièce jointe	`true`pour explorer des fichiers dans vos espaces personnels, pages, blogs, commentaires de page, pièces jointes de page, commentaires de blog et pièces jointes de blog dans Confluence.

Configuration	Description
maxFileSizeInMegaBytes	Spécifiez la limite de taille de fichier MBs Amazon Kendra pouvant être explorée. Amazon Kendra analyse uniquement les fichiers dans la limite de taille que vous avez définie. La taille de fichier par défaut est de 50 Mo. La taille maximale du fichier doit être supérieure à 0 Mo et inférieure ou égale à 50 Mo.
type	Type de source de données. Spécifiez `CONFLUENCEV2` comme type de source de données.
enableIdentityCrawler	`true`utiliser Amazon Kendra le moteur de recherche d'identité pour synchroniser les identity/principal informations relatives aux utilisateurs et aux groupes ayant accès à certains documents. Si le robot d'identification est désactivé, tous les documents peuvent être consultés publiquement. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'identification est désactivé , vous pouvez également utiliser l'PutPrincipalMappingAPI pour télécharger les informations d'accès des utilisateurs et des groupes.

Configuration	Description
Mode de synchronisation	Spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Vous pouvez choisir entre : • `FORCED_FULL_CRAWL` pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index. • `FULL_CRAWL` pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modificat ions de contenu et indexer le contenu modifié depuis la dernière synchronisation.
Secrétaire Arn	Le nom de ressource Amazon (ARN) d'un AWS Secrets Manager secret qui contient les paires clé-valeur requises pour se connecter à votre Confluence. Pour plus d'informations sur ces paires clé-valeur, consultez les instructions de connexion pour Confluence.
version	Version de ce modèle actuellement prise en charge.

Schéma JSON Confluence

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
```

```
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "hostUrl": {
              "type": "string",
              "pattern": "https:.*"
            },
            "type": {
              "type": "string",
              "enum": [
                "SAAS",
                "ON_PREM"
              ]
            },
            "authType": {
              "type": "string",
              "enum": [
                "Basic",
                "OAuth2",
                "Personal-token"
              ]
            }
          },
          "required": [
            "hostUrl",
            "type",
            "authType"
          ]
        }
      },
      "required": [
        "repositoryEndpointMetadata"
      ]
    },
    "repositoryConfigurations": {
      "type": "object",
      "properties": {
        "space": {
          "type": "object",
          "properties": {
            "fieldMappings": {
```

```
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string",
                    "enum": [
                      "STRING",
                      "STRING_LIST",
                      "DATE"
                    ]
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  },
                  "dateFieldFormat": {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required": [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      },
      "page": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
```

```
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string",
                    "enum": [
                      "STRING",
                      "STRING_LIST",
                      "DATE",
                      "LONG"
                    ]
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  },
                  "dateFieldFormat": {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required": [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      },
      "blog": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
```

```
                "indexFieldName": {
                  "type": "string"
                },
                "indexFieldType": {
                  "type": "string",
                  "enum": [
                    "STRING",
                    "STRING_LIST",
                    "DATE",
                    "LONG"
                  ]
                },
                "dataSourceFieldName": {
                  "type": "string"
                },
                "dateFieldFormat": {
                  "type": "string",
                  "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                }
              },
              "required": [
                "indexFieldName",
                "indexFieldType",
                "dataSourceFieldName"
              ]
            }
          ]
        }
      },
      "required": [
        "fieldMappings"
      ]
    },
    "comment": {
      "type": "object",
      "properties": {
        "fieldMappings": {
          "type": "array",
          "items": [
            {
              "type": "object",
              "properties": {
                "indexFieldName": {
                  "type": "string"
```

```
                },
                "indexFieldType": {
                  "type": "string",
                  "enum": [
                    "STRING",
                    "STRING_LIST",
                    "DATE",
                    "LONG"
                  ]
                },
                "dataSourceFieldName": {
                  "type": "string"
                },
                "dateFieldFormat": {
                  "type": "string",
                  "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                }
              },
              "required": [
                "indexFieldName",
                "indexFieldType",
                "dataSourceFieldName"
              ]
            }
          ]
        }
      },
      "required": [
        "fieldMappings"
      ]
    },
    "attachment": {
      "type": "object",
      "properties": {
        "fieldMappings": {
          "type": "array",
          "items": [
            {
              "type": "object",
              "properties": {
                "indexFieldName": {
                  "type": "string"
                },
                "indexFieldType": {
```

```
                    "type": "string",
                    "enum": [
                      "STRING",
                      "STRING_LIST",
                      "DATE",
                      "LONG"
                    ]
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  },
                  "dateFieldFormat": {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required": [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      }
    },
    "additionalProperties": {
      "type": "object",
      "properties": {
        "usersAclS3FilePath": {
          "type": "string"
        },
        "isCrawlAcl": {
          "type": "boolean"
        },
        "fieldForUserId": {
          "type": "string"
        },
        "inclusionSpaceKeyFilter": {
```

```
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "exclusionSpaceKeyFilter": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "pageTitleRegEX": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "blogTitleRegEX": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "commentTitleRegEX": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "attachmentTitleRegEX": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "isCrawlPersonalSpace": {
          "type": "boolean"
        },
        "isCrawlArchivedSpace": {
          "type": "boolean"
        },
        "isCrawlArchivedPage": {
          "type": "boolean"
        },
```

```
        "isCrawlPage": {
          "type": "boolean"
        },
        "isCrawlBlog": {
          "type": "boolean"
        },
        "isCrawlPageComment": {
          "type": "boolean"
        },
        "isCrawlPageAttachment": {
          "type": "boolean"
        },
        "isCrawlBlogComment": {
          "type": "boolean"
        },
        "isCrawlBlogAttachment": {
          "type": "boolean"
        },
        "maxFileSizeInMegaBytes":  {
          "type":"string"
        },
        "inclusionFileTypePatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "exclusionFileTypePatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "inclusionUrlPatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "exclusionUrlPatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
```

```
        },
        "proxyHost": {
          "type": "string"
        },
        "proxyPort": {
          "type": "string"
        }
      },
      "required": []
    },
    "type": {
      "type": "string",
      "pattern": "CONFLUENCEV2"
    },
    "enableIdentityCrawler": {
      "type": "boolean"
    },
    "syncMode": {
      "type": "string",
      "enum": [
        "FULL_CRAWL",
        "FORCED_FULL_CRAWL"
      ]
    },
    "secretArn": {
      "type": "string",
      "minLength": 20,
      "maxLength": 2048
    }
  },
  "version": {
    "type": "string",
    "anyOf": [
      {
        "pattern": "1.0.0"
      }
    ]
  },
  "required": [
    "connectionConfiguration",
    "repositoryConfigurations",
    "syncMode",
    "additionalProperties",
    "secretArn",
```

```
    "type"
  ]
}
```

Schéma du modèle Dropbox

Vous incluez un JSON qui contient le schéma de la source de données dans le cadre de l'TemplateConfigurationobjet. Vous fournissez la clé d'application Dropbox, le secret de l'application et le jeton d'accès dans le cadre du secret qui stocke vos informations d'authentification. Spécifiez également le type de source de donnéesDROPBOX, le type de jeton d'accès que vous souhaitez utiliser (temporaire ou permanent) et les autres configurations nécessaires. Vous spécifiez ensuite `TEMPLATE` le `Type` moment de l'appel CreateDataSource.

Vous pouvez utiliser le modèle fourni dans ce guide du développeur. Consultez Schéma JSON Dropbox.

Le tableau suivant décrit les paramètres du schéma Dropbox JSON.

Configuration	Description
Configuration de connexion	Informations de configuration pour le point de terminaison de la source de données.
repositoryEndpointMetadata	Informations sur le point de terminaison de la source de données. Cette source de données ne spécifie pas de point de terminaison dans `repositoryEndpointMetadata` . Les informations de connexion sont plutôt incluses dans un AWS Secrets Manager secret que vous fournissez `secretArn` .
Configurations du référentiel	Informations de configuration pour le contenu de la source de données. Par exemple, configurer des types spécifiques de contenu et des mappages de champs.
• dans le fichier • paper • papert	Liste d'objets qui mappent les attributs ou les noms de champs de vos fichiers Dropbox, de Dropbox Paper, ainsi que des raccourcis

Configuration	Description
• raccourci	permettant d' Amazon Kendra indexer les noms de champs. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.
Mode de synchronisation	Spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Vous pouvez choisir entre : • `FORCED_FULL_CRAWL` pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index. • `FULL_CRAWL` pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modificat ions de contenu et indexer le contenu modifié depuis la dernière synchronisation. • `CHANGE_LOG` pour indexer uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchroni se avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

Configuration	Description
enableIdentityCrawler	`true`utiliser Amazon Kendra le moteur de recherche d'identité pour synchroniser les identity/principal informations relatives aux utilisateurs et aux groupes ayant accès à certains documents. Si le robot d'identification est désactivé, tous les documents peuvent être consultés publiquement. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'identification est désactivé , vous pouvez également utiliser l'PutPrincipalMappingAPI pour télécharger les informations d'accès des utilisateurs et des groupes.
Secrétaire Arn	Le nom de ressource Amazon (ARN) d'un AWS Secrets Manager secret qui contient les paires clé-valeur requises pour se connecter à votre Dropbox. Le secret doit contenir une structure JSON avec les clés suivantes : `{ "appKey": "Dropbox app key", "appSecret": " Dropbox app secret", "accesstoken": " temporary access token or refresh access token" }`
Propriétés supplémentaires	Options de configuration supplémentaires pour votre contenu dans votre source de données.

Configuration	Description
isCrawlAcl	`true`pour analyser les informations de la liste de contrôle d'accès (ACL) de vos documents , si vous disposez d'une ACL et que vous souhaitez l'utiliser pour le contrôle d'accès. L'ACL indique les documents auxquels les utilisateurs et les groupes peuvent accéder. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.
• inclusionFileNameMotifs • inclusionFileTypeMotifs	Liste de modèles d'expressions régulières permettant d'inclure certains noms et types de fichiers dans votre source de données Dropbox. Les fichiers qui correspondent aux modèles sont inclus dans l'index. Les fichiers qui ne correspondent pas aux modèles sont exclus de l'index. Si un fichier correspond à la fois à un modèle d'inclusion et d'exclusion, le modèle d'exclusion a la priorité et le fichier n'est pas inclus dans l'index.
• exclusionFileNameMotifs • exclusionFileTypeMotifs	Liste de modèles d'expressions régulières permettant d'exclure certains noms et types de fichiers de votre source de données Dropbox. Les fichiers qui correspondent aux modèles sont exclus de l'index. Les fichiers qui ne correspondent pas aux modèles sont inclus dans l'index. Si un fichier correspond à la fois à un modèle d'exclusion et à un modèle d'inclusi on, le modèle d'exclusion est prioritaire et le fichier n'est pas inclus dans l'index.

Configuration	Description
• Fichier d'exploration • Papier Crawl • Crawl Paper • Raccourci Crawl	`true`pour explorer les fichiers de votre Dropbox, les documents Dropbox Paper, les modèles Dropbox Paper et les raccourcis de pages Web stockés dans votre Dropbox.
type	Type de source de données. Spécifiez `DROPBOX` comme type de source de données.
Type de jeton	Spécifiez le type de jeton d'accès : jeton d'accès permanent ou temporaire. Il est recommandé de créer un jeton d'accès d'actualisation qui n'expire jamais dans Dropbox plutôt que de vous fier à un jeton d'accès unique expirant au bout de 4 heures. Vous créez une application et un jeton d'accès actualisé dans la console de développement Dropbox, puis vous fournissez le jeton d'accès dans votre code secret.
version	Version de ce modèle actuellement prise en charge.

Schéma JSON Dropbox

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
          }
        }
```

```
      },
      "required": [
        "repositoryEndpointMetadata"
      ]
    },
    "repositoryConfigurations": {
      "type": "object",
      "properties": {
        "file": {
          "type": "object",
          "properties": {
            "fieldMappings": {
              "type": "array",
              "items": {
                "anyOf": [
                  {
                    "type": "object",
                    "properties": {
                      "indexFieldName": {
                        "type": "string"
                      },
                      "indexFieldType": {
                        "type": "string",
                        "enum": [
                          "STRING",
                          "STRING_LIST",
                          "LONG",
                          "DATE"
                        ]
                      },
                      "dataSourceFieldName": {
                        "type": "string"
                      },
                      "dateFieldFormat": {
                        "type": "string",
                        "pattern": "dd-MM-yyyy HH:mm:ss"
                      }
                    },
                    "required": [
                      "indexFieldName",
                      "indexFieldType",
                      "dataSourceFieldName"
                    ]
                  }
```

```
            ]
          }
        }
      },
      "required": [
        "fieldMappings"
      ]
    },
    "paper": {
      "type": "object",
      "properties": {
        "fieldMappings": {
          "type": "array",
          "items": {
            "anyOf": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string",
                    "enum": [
                      "STRING",
                      "STRING_LIST",
                      "LONG",
                      "DATE"
                    ]
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  },
                  "dateFieldFormat": {
                    "type": "string",
                    "pattern": "dd-MM-yyyy HH:mm:ss"
                  }
                },
                "required": [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
```

```
            ]
          }
        }
      },
      "required": [
        "fieldMappings"
      ]
    },
    "papert": {
      "type": "object",
      "properties": {
        "fieldMappings": {
          "type": "array",
          "items": {
            "anyOf": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string",
                    "enum": [
                      "STRING",
                      "STRING_LIST",
                      "LONG",
                      "DATE"
                    ]
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  },
                  "dateFieldFormat": {
                    "type": "string",
                    "pattern": "dd-MM-yyyy HH:mm:ss"
                  }
                },
                "required": [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
```

```
            ]
          }
        }
      },
      "required": [
        "fieldMappings"
      ]
    },
    "shortcut": {
      "type": "object",
      "properties": {
        "fieldMappings": {
          "type": "array",
          "items": {
            "anyOf": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string",
                    "enum": [
                      "STRING",
                      "STRING_LIST",
                      "LONG",
                      "DATE"
                    ]
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  },
                  "dateFieldFormat": {
                    "type": "string",
                    "pattern": "dd-MM-yyyy HH:mm:ss"
                  }
                },
                "required": [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
```

```
              ]
            }
          }
        },
        "required": [
          "fieldMappings"
        ]
      }
    }
  },
  "syncMode": {
    "type": "string",
    "enum": [
      "FULL_CRAWL",
      "FORCED_FULL_CRAWL",
      "CHANGE_LOG"
    ]
  },
  "enableIdentityCrawler": {
    "type": "boolean"
  },
  "secretArn": {
    "type": "string"
  },
  "additionalProperties": {
    "type": "object",
    "properties": {
      "isCrawlAcl": {
        "type": "boolean"
      },
      "inclusionFileNamePatterns": {
        "type": "array"
      },
      "exclusionFileNamePatterns": {
        "type": "array"
      },
      "inclusionFileTypePatterns": {
        "type": "array"
      },
      "exclusionFileTypePatterns": {
        "type": "array"
      },
      "crawlFile": {
        "type": "boolean"
```

```
        },
        "crawlPaper": {
          "type": "boolean"
        },
        "crawlPapert": {
          "type": "boolean"
        },
        "crawlShortcut": {
          "type": "boolean"
        }
      }
    },
    "type": {
      "type": "string",
      "pattern": "DROPBOX"
    },
    "tokenType": {
      "type": "string",
      "enum": [
        "PERMANENT",
        "TEMPORARY"
      ]
    },
    "version": {
      "type": "string",
      "anyOf": [
        {
          "pattern": "1.0.0"
        }
      ]
    }
  },
  "additionalProperties": false,
  "required": [
    "connectionConfiguration",
    "repositoryConfigurations",
    "additionalProperties",
    "syncMode",
    "enableIdentityCrawler",
    "secretArn",
    "type",
    "tokenType"
  ]
```

```
}
```

Schéma du modèle Drupal

Vous incluez un JSON qui contient le schéma de la source de données dans le cadre de l'TemplateConfigurationobjet. Vous fournissez l'URL de l'hôte Drupal et le type d'authentification dans le cadre de la configuration de la connexion ou des détails du point de terminaison du référentiel. Spécifiez également le type de source de données DRUPAL, un secret pour vos informations d'authentification et les autres configurations nécessaires. Vous spécifiez ensuite `TEMPLATE` le `Type` moment de l'appel CreateDataSource.

Vous pouvez utiliser le modèle fourni dans ce guide du développeur. Consultez Schéma JSON Drupal.

Le tableau suivant décrit les paramètres du schéma JSON de Drupal.

Configuration	Description
Configuration de connexion	Informations de configuration pour le point de terminaison de la source de données.
repositoryEndpointMetadata	Informations sur le point de terminaison de la source de données.
URL de l'hôte	L'URL de l'hôte de votre site Web Drupal. Par exemple, *`https://<hostname>/<drupalsitename>`* .
Configurations du référentiel	Informations de configuration pour le contenu de la source de données.
• content • comment • attachment	Une liste d'objets qui mappent les attributs ou les noms de champs de vos fichiers Drupal. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données. Les noms des champs de source de données Drupal doivent exister dans vos métadonnées personnalisées Drupal.

Configuration	Description
Propriétés supplémentaires	Options de configuration supplémentaires pour votre contenu dans votre source de données.
• inclusionFileNameMotifs • articleTitleInclusionMotifs • pageTitleInclusionMotifs • customContentTitleInclusionPatterns • basicBlockTitleInclusionPatterns • customBlockTitleInclusionPatterns	Une liste de modèles d'expressions régulières pour inclure certains fichiers dans votre source de données Drupal. Les fichiers qui correspondent aux modèles sont inclus dans l'index. Les fichiers qui ne correspondent pas aux modèles sont exclus de l'index. Si un fichier correspond à la fois à un modèle d'inclusion et d'exclusion, le modèle d'exclusion a la priorité et le fichier n'est pas inclus dans l'index.
• exclusionFileNameMotifs • articleTitleExclusionMotifs • pageTitleExclusionMotifs • customContentTitleExclusionPatterns • basicBlockTitleExclusionPatterns • customBlockTitleExclusionPatterns	Une liste de modèles d'expressions régulières pour exclure certains fichiers de votre source de données Drupal. Les fichiers qui correspondent aux modèles sont exclus de l'index. Les fichiers qui ne correspondent pas aux modèles sont inclus dans l'index. Si un fichier correspond à la fois à un modèle d'exclusion et à un modèle d'inclusion, le modèle d'exclusion est prioritaire et le fichier n'est pas inclus dans l'index.
Définitions du contenu • contentType • Définition du champ • isCrawlComments • isCrawlFiles • isCrawlArticle • isCrawlBasicPage • isCrawlBasicBloquer • isCrawlCustomContentTypesList	Spécifiez les types de contenu à analyser et indiquez si vous souhaitez analyser les commentaires et les pièces jointes pour les types de contenu que vous avez sélectionnés.

Configuration	Description
type	Type de source de données. Spécifiez `DRUPAL` comme type de source de données.
authType	Le type d'authentification que vous utilisez, que ce soit `BASIC-AUTH` ou `OAUTH2`.
Mode de synchronisation	Spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Vous pouvez choisir entre : • `FORCED_FULL_CRAWL` pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index. • `FULL_CRAWL` pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modificat ions de contenu et indexer le contenu modifié depuis la dernière synchronisation. • `CHANGE_LOG` pour indexer uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchroni se avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

Configuration	Description
enableIdentityCrawler	`true`utiliser Amazon Kendra le moteur de recherche d'identité pour synchroniser les identity/principal informations relatives aux utilisateurs et aux groupes ayant accès à certains documents. Si le robot d'identification est désactivé, tous les documents peuvent être consultés publiquement. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'identification est désactivé , vous pouvez également utiliser l'PutPrincipalMappingAPI pour télécharger les informations d'accès des utilisateurs et des groupes.

Secrétaire Arn	Le nom de ressource Amazon (ARN) d'un AWS Secrets Manager secret qui contient les paires clé-valeur requises pour se connecter à votre Drupal. Le secret doit contenir une structure JSON avec les clés suivantes :
	
	Si vous utilisez l'authentification de base :
	
	`{ "username": "user name", "passwords": "password" }`
	
	Si vous utilisez l'authentification OAuth 2.0 :
	
	`{ "username": "user name", "password": "password" , "clientId": "client id", "clientSecret": "client secret" }`

Configuration	Description
version	Version de ce modèle actuellement prise en charge.

Schéma JSON Drupal

```
{
 "$schema": "http://json-schema.org/draft-04/schema#",
 "type": "object",
 "properties": {
  "connectionConfiguration": {
   "type": "object",
   "properties": {
    "repositoryEndpointMetadata": {
     "type": "object",
     "properties": {
      "hostUrl": {
       "type": "string",
       "pattern": "https:.*"
      }
     },
     "required": [
      "hostUrl"
     ]
    }
   },
   "required": [
    "repositoryEndpointMetadata"
   ]
  },
  "repositoryConfigurations": {
   "type": "object",
   "properties": {
    "content": {
     "type": "object",
     "properties": {
      "fieldMappings": {
       "type": "array",
       "items": [
        {
         "type": "object",
```

```
        "properties": {
         "indexFieldName": {
          "type": "string"
         },
         "indexFieldType": {
          "type": "string",
          "enum": [
           "STRING",
           "DATE"
          ]
         },
         "dataSourceFieldName": {
          "type": "string"
         },
         "dateFieldFormat": {
          "type": "string",
          "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
         }
        },
        "required": [
         "indexFieldName",
         "indexFieldType",
         "dataSourceFieldName"
        ]
       }
      ]
     }
    },
    "required": [
     "fieldMappings"
    ]
   },
   "comment": {
    "type": "object",
    "properties": {
     "fieldMappings": {
      "type": "array",
      "items": [
       {
        "type": "object",
        "properties": {
         "indexFieldName": {
          "type": "string"
         },
```

```
          "indexFieldType": {
           "type": "string",
           "enum": [
            "STRING",
            "DATE"
           ]
          },
          "dataSourceFieldName": {
           "type": "string"
          },
          "dateFieldFormat": {
           "type": "string",
           "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
          }
         },
         "required": [
          "indexFieldName",
          "indexFieldType",
          "dataSourceFieldName"
         ]
        }
       ]
      }
     },
     "required": [
      "fieldMappings"
     ]
    },
    "attachment": {
     "type": "object",
     "properties": {
      "fieldMappings": {
       "type": "array",
       "items": [
        {
         "type": "object",
         "properties": {
          "indexFieldName": {
           "type": "string"
          },
          "indexFieldType": {
           "type": "string",
           "enum": [
            "STRING",
```

```
            "DATE"
           ]
          },
          "dataSourceFieldName": {
           "type": "string"
          },
          "dateFieldFormat": {
           "type": "string",
           "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
          }
         },
         "required": [
          "indexFieldName",
          "indexFieldType",
          "dataSourceFieldName"
         ]
        }
       ]
      }
     },
     "required": [
      "fieldMappings"
     ]
    }
   }
  },
  "additionalProperties": {
   "type": "object",
   "properties": {
    "isCrawlArticle": {
     "type": "boolean"
    },
    "isCrawlBasicPage": {
     "type": "boolean"
    },
    "isCrawlBasicBlock": {
     "type": "boolean"
    },
    "crawlCustomContentTypesList": {
     "type": "array",
     "items": {
      "type": "string"
     }
    },
```

```
    "crawlCustomBlockTypesList": {
     "type": "array",
     "items": {
      "type": "string"
     }
    },
    "filePath": {
     "anyOf": [
      {
       "type": "string",
       "pattern": "s3:.*"
      },
      {
       "type": "string",
       "pattern": ""
      }
     ]
    },
    "inclusionFileNamePatterns": {
     "type": "array",
     "items": {
      "type": "string"
     }
    },
    "exclusionFileNamePatterns": {
     "type": "array",
     "items": {
      "type": "string"
     }
    },
    "articleTitleInclusionPatterns": {
     "type": "array",
     "items": {
      "type": "string"
     }
    },
    "articleTitleExclusionPatterns": {
     "type": "array",
     "items": {
      "type": "string"
     }
    },
    "pageTitleInclusionPatterns": {
     "type": "array",
```

```
      "items": {
       "type": "string"
      }
     },
     "pageTitleExclusionPatterns": {
      "type": "array",
      "items": {
       "type": "string"
      }
     },
     "customContentTitleInclusionPatterns": {
      "type": "array",
      "items": {
       "type": "string"
      }
     },
     "customContentTitleExclusionPatterns": {
      "type": "array",
      "items": {
       "type": "string"
      }
     },
     "basicBlockTitleInclusionPatterns": {
      "type": "array",
      "items": {
       "type": "string"
      }
     },
     "basicBlockTitleExclusionPatterns": {
      "type": "array",
      "items": {
       "type": "string"
      }
     },
     "customBlockTitleInclusionPatterns": {
      "type": "array",
      "items": {
       "type": "string"
      }
     },
     "customBlockTitleExclusionPatterns": {
      "type": "array",
      "items": {
       "type": "string"
```

```
     }
    },
    "contentDefinitions": {
     "type": "array",
     "items": {
      "properties": {
       "contentType": {
        "type": "string"
       },
       "fieldDefinition": {
        "type": "array",
        "items": [
         {
          "type": "object",
          "properties": {
           "machineName": {
            "type": "string"
           },
           "type": {
            "type": "string"
           }
          },
          "required": [
           "machineName",
           "type"
          ]
         }
        ]
       },
       "isCrawlComments": {
        "type": "boolean"
       },
       "isCrawlFiles": {
        "type": "boolean"
       }
      }
     },
     "required": [
      "contentType",
      "fieldDefinition",
      "isCrawlComments",
      "isCrawlFiles"
     ]
    }
```

```
    },
    "required": []
   },
   "type": {
    "type": "string",
    "pattern": "DRUPAL"
   },
   "authType": {
    "type": "string",
    "enum": [
     "BASIC-AUTH",
     "OAUTH2"
    ]
   },
   "syncMode": {
    "type": "string",
    "enum": [
     "FORCED_FULL_CRAWL",
     "FULL_CRAWL",
     "CHANGE_LOG"
    ]
   },
   "enableIdentityCrawler": {
    "type": "boolean"
   },
   "secretArn": {
    "type": "string",
    "minLength": 20,
    "maxLength": 2048
   }
  },
  "version": {
   "type": "string",
   "anyOf": [
    {
     "pattern": "1.0.0"
    }
   ]
  },
  "required": [
   "connectionConfiguration",
   "repositoryConfigurations",
   "syncMode",
   "additionalProperties",
```

```
  "secretArn",
  "type"
 ]
}
```

GitHub schéma de modèle

Vous incluez un JSON qui contient le schéma de la source de données dans le cadre de l'TemplateConfigurationobjet. Vous indiquez l'URL de l' GitHub hôte, le nom de l'organisation et indiquez si vous utilisez le GitHub GitHub cloud ou sur site dans le cadre de la configuration de la connexion ou des détails du point de terminaison du référentiel. Spécifiez également le type de source de donnéesGITHUB, un secret pour vos informations d'authentification et les autres configurations nécessaires. Vous spécifiez ensuite TEMPLATE le Type moment de l'appel CreateDataSource.

Vous pouvez utiliser le modèle fourni dans ce guide du développeur. Consultez GitHub Schéma JSON.

Le tableau suivant décrit les paramètres du schéma GitHub JSON.

Configuration	Description
Configuration de connexion	Informations de configuration pour le point de terminaison de la source de données.
repositoryEndpointMetadata	Informations sur le point de terminaison de la source de données.
type	Spécifiez le type sous la forme SAAS ouON_PREMISE .
URL de l'hôte	URL de l' GitHub hôte. Par exemple, si vous utilisez le GitHub SaaS/Enterprise Cloud :. https://api.github.com Ou, si vous utilisez un serveur GitHub local ou d'entreprise :. https://on-prem-host-url/api/v3/
Nom de l'organisation	Vous pouvez trouver le nom de votre organisation lorsque vous vous connectez à votre GitHub ordinateur de bureau et que vous

Configuration	Description
	accédez à Vos organisations dans le menu déroulant de votre photo de profil.
Configurations du référentiel	Informations de configuration pour le contenu de la source de données. Par exemple, configurer des types spécifiques de contenu et des mappages de champs.
• Référentiel GH • GHCommit • ghIssueDocument • ghIssueComment • ghIssueAttachment • gh PRDocument • gh PRComment • gh PRAttachment	Liste d'objets qui mappent les attributs ou les noms de champs de votre GitHub contenu aux noms de champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.
Propriétés supplémentaires	Options de configuration supplémentaires pour votre contenu dans votre source de données.
isCrawlAcl	`true`pour analyser les informations de la liste de contrôle d'accès (ACL) de vos documents , si vous disposez d'une ACL et que vous souhaitez l'utiliser pour le contrôle d'accès. L'ACL indique les documents auxquels les utilisateurs et les groupes peuvent accéder et effectuer des recherches. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.

Configuration	Description
fieldForUserID	Spécifiez le type d'ID utilisateur que vous souhaitez utiliser pour l'analyse des ACL. Spécifiez `email` si vous souhaitez utiliser l'adresse e-mail de l'utilisateur pour l'ID utilisateur ou `username` si vous souhaitez utiliser le nom d'utilisateur pour l'ID utilisateur. Si vous ne spécifiez aucune option, elle `email` est utilisée par défaut.
Filtre de référentiel	Liste des noms des référentiels et des noms de branches spécifiques que vous souhaitez indexer.
Référentiel Crawl	`true`pour explorer les référentiels.
crawlRepositoryDocuments	`true`pour explorer les documents du référentiel.
Problème Crawlish	`true`pour les problèmes de crawl.
crawlIssueComment	`true`pour explorer les commentaires des problèmes.
crawlIssueCommentPièce jointe	`true`pour explorer les pièces jointes aux commentaires des problèmes.
crawlPullRequest	`true`pour explorer les pull requests.
crawlPullRequestCommentaire	`true`pour explorer les commentaires des pull requests.
crawlPullRequestCommentAttachment	`true`pour explorer les pièces jointes aux commentaires de la pull request.

Configuration	Description
• inclusionFolderNameMotifs • inclusionFileTypeMotifs • inclusionFileNameMotifs	Liste de modèles d'expressions régulières permettant d'inclure certains contenus dans votre source de GitHub données. Le contenu qui correspond aux modèles est inclus dans l'index. Le contenu qui ne correspond pas aux modèles est exclu de l'index. Si un contenu correspond à la fois à un modèle d'inclusion et à un modèle d'exclusion, le modèle d'exclusi on est prioritaire et le contenu n'est pas inclus dans l'index.
• exclusionFolderNameMotifs • exclusionFileTypeMotifs • exclusionFileNameMotifs	Liste de modèles d'expressions régulières permettant d'exclure certains contenus de votre source de GitHub données. Le contenu qui correspond aux modèles est exclu de l'index. Le contenu qui ne correspond pas aux modèles est inclus dans l'index. Si un contenu correspon d à la fois à un modèle d'inclusion et à un modèle d'exclusion, le modèle d'exclusion est prioritaire et le contenu n'est pas inclus dans l'index.
type	Type de source de données. Spécifiez `GITHUB` comme type de source de données.

Configuration	Description
enableIdentityCrawler	`true`utiliser Amazon Kendra le moteur de recherche d'identité pour synchroniser les identity/principal informations relatives aux utilisateurs et aux groupes ayant accès à certains documents. Si le robot d'identification est désactivé, tous les documents peuvent être consultés publiquement. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'identification est désactivé , vous pouvez également utiliser l'PutPrincipalMappingAPI pour télécharger les informations d'accès des utilisateurs et des groupes.

Configuration	Description
Mode de synchronisation	Spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Vous pouvez choisir entre : • `FORCED_FULL_CRAWL` pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index. • `FULL_CRAWL` pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation. • `CHANGE_LOG` pour indexer uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

Configuration	Description
Secrétaire Arn	Le nom de ressource Amazon (ARN) d'un AWS Secrets Manager secret qui contient les paires clé-valeur requises pour se connecter à votre. GitHub Le secret doit contenir une structure JSON avec les clés suivantes : `{ "personalToken": " token" }`
version	Version de ce modèle actuellement prise en charge.

GitHub Schéma JSON

Le schéma GitHub JSON est le suivant :

```
{
    "$schema": "http://json-schema.org/draft-04/schema#",
    "type": "object",
    "properties": {
        "connectionConfiguration": {
            "type": "object",
            "properties": {
                "repositoryEndpointMetadata": {
                    "type": "object",
                    "properties": {
                        "type": {
                            "type": "string"
                        },
                        "hostUrl": {
                            "type": "string",
                            "pattern": "https://.*"
                        },
                        "organizationName": {
                            "type": "string"
                        }
                    },
```

```
                "required": [
                    "type",
                    "hostUrl",
                    "organizationName"
                ]
            }
        },
        "required": [
            "repositoryEndpointMetadata"
        ]
    },
    "repositoryConfigurations": {
        "type": "object",
        "properties": {
            "ghRepository": {
                "type": "object",
                "properties": {
                    "fieldMappings": {
                        "type": "array",
                        "items": [
                            {
                                "type": "object",
                                "properties": {
                                    "indexFieldName": {
                                        "type": "string"
                                    },
                                    "indexFieldType": {
                                        "type": "string",
                                        "enum": [
                                            "STRING",
                                            "STRING_LIST",
                                            "DATE"
                                        ]
                                    },
                                    "dataSourceFieldName": {
                                        "type": "string"
                                    },
                                    "dateFieldFormat": {
                                        "type": "string",
                                        "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                                    }
                                },
                                "required": [
                                    "indexFieldName",
```

```
                                "indexFieldType",
                                "dataSourceFieldName"
                            ]
                        }
                    ]
                }
            },
            "required": [
                "fieldMappings"
            ]
        },
        "ghCommit": {
            "type": "object",
            "properties": {
                "fieldMappings": {
                    "type": "array",
                    "items": [
                        {
                            "type": "object",
                            "properties": {
                                "indexFieldName": {
                                    "type": "string"
                                },
                                "indexFieldType": {
                                    "type": "string",
                                    "enum": [
                                        "STRING",
                                        "STRING_LIST",
                                        "DATE"
                                    ]
                                },
                                "dataSourceFieldName": {
                                    "type": "string"
                                },
                                "dateFieldFormat": {
                                    "type": "string",
                                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                                }
                            },
                            "required": [
                                "indexFieldName",
                                "indexFieldType",
                                "dataSourceFieldName"
                            ]
```

```
                        }
                    ]
                }
            },
            "required": [
                "fieldMappings"
            ]
        },
        "ghIssueDocument": {
            "type": "object",
            "properties": {
                "fieldMappings": {
                    "type": "array",
                    "items": [
                        {
                            "type": "object",
                            "properties": {
                                "indexFieldName": {
                                    "type": "string"
                                },
                                "indexFieldType": {
                                    "type": "string",
                                    "enum": [
                                        "STRING",
                                        "STRING_LIST",
                                        "DATE"
                                    ]
                                },
                                "dataSourceFieldName": {
                                    "type": "string"
                                },
                                "dateFieldFormat": {
                                    "type": "string",
                                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                                }
                            },
                            "required": [
                                "indexFieldName",
                                "indexFieldType",
                                "dataSourceFieldName"
                            ]
                        }
                    ]
                }
```

```
                    },
                    "required": [
                        "fieldMappings"
                    ]
                },
                "ghIssueComment": {
                    "type": "object",
                    "properties": {
                        "fieldMappings": {
                            "type": "array",
                            "items": [
                                {
                                    "type": "object",
                                    "properties": {
                                        "indexFieldName": {
                                            "type": "string"
                                        },
                                        "indexFieldType": {
                                            "type": "string",
                                            "enum": [
                                                "STRING",
                                                "STRING_LIST",
                                                "DATE"
                                            ]
                                        },
                                        "dataSourceFieldName": {
                                            "type": "string"
                                        },
                                        "dateFieldFormat": {
                                            "type": "string",
                                            "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                                        }
                                    },
                                    "required": [
                                        "indexFieldName",
                                        "indexFieldType",
                                        "dataSourceFieldName"
                                    ]
                                }
                            ]
                        }
                    },
                    "required": [
                        "fieldMappings"
```

```
                ]
            },
            "ghIssueAttachment": {
                "type": "object",
                "properties": {
                    "fieldMappings": {
                        "type": "array",
                        "items": [
                            {
                                "type": "object",
                                "properties": {
                                    "indexFieldName": {
                                        "type": "string"
                                    },
                                    "indexFieldType": {
                                        "type": "string",
                                        "enum": [
                                            "STRING",
                                            "STRING_LIST",
                                            "DATE"
                                        ]
                                    },
                                    "dataSourceFieldName": {
                                        "type": "string"
                                    },
                                    "dateFieldFormat": {
                                        "type": "string",
                                        "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                                    }
                                },
                                "required": [
                                    "indexFieldName",
                                    "indexFieldType",
                                    "dataSourceFieldName"
                                ]
                            }
                        ]
                    }
                },
                "required": [
                    "fieldMappings"
                ]
            },
            "ghPRDocument": {
```

```
                    "type": "object",
                    "properties": {
                        "fieldMappings": {
                            "type": "array",
                            "items": [
                                {
                                    "type": "object",
                                    "properties": {
                                        "indexFieldName": {
                                            "type": "string"
                                        },
                                        "indexFieldType": {
                                            "type": "string",
                                            "enum": [
                                                "STRING",
                                                "STRING_LIST",
                                                "DATE"
                                            ]
                                        },
                                        "dataSourceFieldName": {
                                            "type": "string"
                                        },
                                        "dateFieldFormat": {
                                            "type": "string",
                                            "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                                        }
                                    },
                                    "required": [
                                        "indexFieldName",
                                        "indexFieldType",
                                        "dataSourceFieldName"
                                    ]
                                }
                            ]
                        }
                    },
                    "required": [
                        "fieldMappings"
                    ]
                },
                "ghPRComment": {
                    "type": "object",
                    "properties": {
                        "fieldMappings": {
```

```
                        "type": "array",
                        "items": [
                            {
                                "type": "object",
                                "properties": {
                                    "indexFieldName": {
                                        "type": "string"
                                    },
                                    "indexFieldType": {
                                        "type": "string",
                                        "enum": [
                                            "STRING",
                                            "STRING_LIST",
                                            "DATE"
                                        ]
                                    },
                                    "dataSourceFieldName": {
                                        "type": "string"
                                    },
                                    "dateFieldFormat": {
                                        "type": "string",
                                        "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                                    }
                                },
                                "required": [
                                    "indexFieldName",
                                    "indexFieldType",
                                    "dataSourceFieldName"
                                ]
                            }
                        ]
                    }
                },
                "required": [
                    "fieldMappings"
                ]
            },
            "ghPRAttachment": {
                "type": "object",
                "properties": {
                    "fieldMappings": {
                        "type": "array",
                        "items": [
                            {
```

```
                                "type": "object",
                                "properties": {
                                    "indexFieldName": {
                                        "type": "string"
                                    },
                                    "indexFieldType": {
                                        "type": "string",
                                        "enum": [
                                            "STRING",
                                            "STRING_LIST",
                                            "DATE"
                                        ]
                                    },
                                    "dataSourceFieldName": {
                                        "type": "string"
                                    },
                                    "dateFieldFormat": {
                                        "type": "string",
                                        "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                                    }
                                },
                                "required": [
                                    "indexFieldName",
                                    "indexFieldType",
                                    "dataSourceFieldName"
                                ]
                            }
                        ]
                    }
                },
                "required": [
                    "fieldMappings"
                ]
            }
        }
    },
    "additionalProperties": {
        "type": "object",
        "properties": {
            "isCrawlAcl": {
                "type": "boolean"
            },
            "fieldForUserId": {
                "type": "string"
```

```
                },
                "crawlRepository": {
                    "type": "boolean"
                },
                "crawlRepositoryDocuments": {
                    "type": "boolean"
                },
                "crawlIssue": {
                    "type": "boolean"
                },
                "crawlIssueComment": {
                    "type": "boolean"
                },
                "crawlIssueCommentAttachment": {
                    "type": "boolean"
                },
                "crawlPullRequest": {
                    "type": "boolean"
                },
                "crawlPullRequestComment": {
                    "type": "boolean"
                },
                "crawlPullRequestCommentAttachment": {
                    "type": "boolean"
                },
                "repositoryFilter": {
                    "type": "array",
                    "items": [
                        {
                            "type": "object",
                            "properties": {
                                "repositoryName": {
                                    "type": "string"
                                },
                                "branchNameList": {
                                    "type": "array",
                                    "items": {
                                        "type": "string"
                                    }
                                }
                            }
                        }
                    ]
                },
```

```
                "inclusionFolderNamePatterns": {
                    "type": "array",
                    "items": {
                        "type": "string"
                    }
                },
                "inclusionFileTypePatterns": {
                    "type": "array",
                    "items": {
                        "type": "string"
                    }
                },
                "inclusionFileNamePatterns": {
                    "type": "array",
                    "items": {
                        "type": "string"
                    }
                },
                "exclusionFolderNamePatterns": {
                    "type": "array",
                    "items": {
                        "type": "string"
                    }
                },
                "exclusionFileTypePatterns": {
                    "type": "array",
                    "items": {
                        "type": "string"
                    }
                },
                "exclusionFileNamePatterns": {
                    "type": "array",
                    "items": {
                        "type": "string"
                    }
                }
            },
            "required": []
        },
        "type": {
            "type": "string",
            "pattern": "GITHUB"
        },
        "syncMode": {
```

```
            "type": "string",
            "enum": [
                "FULL_CRAWL",
                "FORCED_FULL_CRAWL",
                "CHANGE_LOG"
            ]
        },
        "enableIdentityCrawler": {
            "type": "boolean"
        },
        "secretArn": {
            "type": "string",
            "minLength": 20,
            "maxLength": 2048
        }
    },
    "version": {
        "type": "string",
        "anyOf": [
            {
                "pattern": "1.0.0"
            }
        ]
    },
    "required": [
        "connectionConfiguration",
        "repositoryConfigurations",
        "syncMode",
        "additionalProperties",
        "enableIdentityCrawler"
    ]
}
```

Schéma du modèle Gmail

Vous incluez un JSON qui contient le schéma de la source de données dans le cadre de l'TemplateConfigurationobjet. Spécifiez le type de source de donnéesGMAIL, un secret pour vos informations d'authentification et les autres configurations nécessaires. Vous spécifiez ensuite TEMPLATE le Type moment de l'appel CreateDataSource.

Vous pouvez utiliser le modèle fourni dans ce guide du développeur. Consultez Schéma JSON Gmail.

Le tableau suivant décrit les paramètres du schéma JSON Gmail.

Configuration	Description
Configuration de connexion	Informations de configuration pour le point de terminaison de la source de données.
repositoryEndpointMetadata	Informations sur le point de terminaison de la source de données. Cette source de données ne spécifie pas de point de terminaison dans `repositoryEndpointMetadata` . Les informations de connexion sont plutôt incluses dans un AWS Secrets Manager secret que vous fournissez `secretArn` .
Configurations du référentiel	Informations de configuration pour le contenu de la source de données. Par exemple, configurer des types spécifiques de contenu et des mappages de champs. Spécifiez le type de source de données et l'ARN secret.
• message • pièces jointes	Liste d'objets qui mappent les attributs ou les noms de champs de vos messages Gmail et de vos pièces jointes aux noms de champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.
Propriétés supplémentaires	Options de configuration supplémentaires pour votre contenu dans votre source de données.
• inclusionLabelNameMotifs • exclusionLabelNameMotifs • inclusionAttachmentTypeMotifs • exclusionAttachmentTypeMotifs • inclusionAttachmentNameMotifs • exclusionAttachmentNameMotifs	Liste de modèles d'expressions régulières permettant d'inclure ou d'exclure des messages portant des noms d'objet spécifiques dans votre source de données Gmail. Les fichiers qui correspondent aux modèles sont inclus dans l'index. Si un fichier correspond à la fois à un modèle d'inclusion et à un modèle d'exclusion,

Configuration	Description
• inclusionSubjectFilter • exclusionSubjectFilter • isSubjectAnd • inclusionFromFilter • exclusionFromFilter • inclusionToFilter • exclusionToFilter • inclusionCcFilter • exclusionCcFilter • inclusionBccFilter • exclusionBccFilter	le modèle d'exclusion est prioritaire et le fichier n'est pas inclus dans l'index.
beforeDateFilter	Spécifiez les messages et les pièces jointes à inclure avant une certaine date.
afterDateFilter	Spécifiez les messages et les pièces jointes à inclure après une certaine date.
isCrawlAttachment	Valeur booléenne permettant de choisir si vous souhaitez analyser les pièces jointes. Les messages sont automatiquement analysés.
type	Type de source de données. Spécifiez `GMAIL` comme type de source de données.
shouldCrawlDraftMessages	Valeur booléenne permettant de choisir si vous souhaitez analyser les brouillons de messages.

Configuration	Description
Mode de synchronisation	Spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Vous pouvez choisir entre : • `FORCED_FULL_CRAWL` pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index. • `FULL_CRAWL` pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation. **Important** Comme il n'existe aucune API permettant de mettre à jour les messages Gmail définitivement supprimés, tout contenu nouveau, modifié ou supprimé est synchronisé : • Ne supprimera pas de votre Amazon Kendra index les messages définitivement supprimés de Gmail • Ne synchronisera pas les modifications dans les libellés des e-mails Gmail

Configuration	Description
	Pour synchroniser les modifications apportées à l'étiquette de votre source de données Gmail et les e-mails définitivement supprimés avec votre Amazon Kendra index, vous devez effectuer régulièrement des analyses complètes.
Secrétaire Arn	Le nom de ressource Amazon (ARN) d'un secret de Secrets Manager qui contient les paires clé-valeur requises pour se connecter à votre compte Gmail. Le secret doit contenir une structure JSON avec les clés suivantes : `{ "adminAccountEmailId": " service account email", "clientEmailId": " user account email", "privateKey": " private key" }`
version	Version du modèle actuellement prise en charge.

Schéma JSON Gmail

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
      }
```

```
    },
    "repositoryConfigurations": {
      "type": "object",
      "properties": {
        "message": {
          "type": "object",
          "properties": {
            "fieldMappings": {
              "type": "array",
              "items": [
                {
                  "type": "object",
                  "properties": {
                    "indexFieldName": {
                      "type": "string"
                    },
                    "indexFieldType": {
                      "type": "string",
                      "enum": ["STRING", "STRING_LIST", "DATE"]
                    },
                    "dataSourceFieldName": {
                      "type": "string"
                    },
                    "dateFieldFormat": {
                      "type": "string"
                    }
                  },
                  "required": [
                    "indexFieldName",
                    "indexFieldType",
                    "dataSourceFieldName"
                  ]
                }
              ]
            }
          }
        },
        "attachments": {
          "type": "object",
          "properties": {
            "fieldMappings": {
              "type": "array",
              "items": [
                {
```

```
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string",
                    "enum": ["STRING"]
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  }
                },
                "required": [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        }
      }
    },
    "required": []
  },
  "additionalProperties": {
    "type": "object",
    "properties": {
      "inclusionLabelNamePatterns": {
        "type": "array",
        "items": {
          "type": "string"
        }
      },
      "exclusionLabelNamePatterns": {
        "type": "array",
        "items": {
          "type": "string"
        }
      },
      "inclusionAttachmentTypePatterns": {
        "type": "array",
        "items": {
```

```
            "type": "string"
          }
        },
        "exclusionAttachmentTypePatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "inclusionAttachmentNamePatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "exclusionAttachmentNamePatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "inclusionSubjectFilter": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "exclusionSubjectFilter": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "isSubjectAnd": {
          "type": "boolean"
        },
        "inclusionFromFilter": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "exclusionFromFilter": {
          "type": "array",
```

```
          "items": {
            "type": "string"
          }
        },
        "inclusionToFilter": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "exclusionToFilter": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "inclusionCcFilter": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "exclusionCcFilter": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "inclusionBccFilter": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "exclusionBccFilter": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "beforeDateFilter": {
          "anyOf": [
            {
              "type": "string",
```

```
            "pattern": "^[0-9]{4}-[0-9]{2}-[0-9]{2}T[0-9]{2}:[0-9]{2}:[0-9]{2}Z$"
          },
          {
            "type": "string",
            "pattern": ""
          }
        ]
      },
      "afterDateFilter": {
        "anyOf": [
          {
            "type": "string",
            "pattern": "^[0-9]{4}-[0-9]{2}-[0-9]{2}T[0-9]{2}:[0-9]{2}:[0-9]{2}Z$"
          },
          {
            "type": "string",
            "pattern": ""
          }
        ]
      },
      "isCrawlAttachment": {
        "type": "boolean"
      },
      "shouldCrawlDraftMessages": {
        "type": "boolean"
      }
    },
    "required": [
      "isCrawlAttachment",
      "shouldCrawlDraftMessages"
    ]
  },
  "type" : {
    "type" : "string",
    "pattern": "GMAIL"
  },
  "syncMode": {
    "type": "string",
    "enum": [
      "FORCED_FULL_CRAWL",
      "FULL_CRAWL"
    ]
  },
  "secretArn": {
```

```
      "type": "string"
    },
    "version": {
      "type": "string",
      "anyOf": [
        {
          "pattern": "1.0.0"
        }
      ]
    }
  },
  "required": [
    "connectionConfiguration",
    "repositoryConfigurations",
    "additionalProperties",
    "syncMode",
    "secretArn",
    "type"
  ]
}
```

Schéma du modèle Google Drive

Vous incluez un JSON qui contient le schéma de la source de données dans le cadre de l'TemplateConfigurationobjet. Spécifiez le type de source de donnéesGOOGLEDRIVE2, un secret pour vos informations d'authentification et les autres configurations nécessaires. Vous spécifiez ensuite TEMPLATE le Type moment de l'appel CreateDataSource.

Vous pouvez utiliser le modèle fourni dans ce guide du développeur. Consultez Schéma JSON de Google Drive.

Le tableau suivant décrit les paramètres du schéma JSON de Google Drive.

Configuration	Description
Configuration de connexion	Informations de configuration pour la source de données.
repositoryEndpointMetadata	Informations sur le point de terminaison de la source de données. Cette source de données ne spécifie pas de point de terminaison.

Configuration	Description
	Vous choisissez votre type d'authentification : `serviceAccount` et`0Auth2`. Les informations de connexion sont incluses dans un AWS Secrets Manager secret que vous fournisse `zsecretArn` .
authType	Choisissez entre `serviceAccount` et `0Auth2` en fonction de votre cas d'utilisation.
Configurations du référentiel	Informations de configuration pour le contenu de la source de données. Par exemple, configurer des types spécifiques de contenu et des mappages de champs.
• dans le fichier • comment	Liste d'objets qui associent les attributs ou les noms de champs de votre Google Drive à des noms de champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.
Propriétés supplémentaires	Options de configuration supplémentaires pour votre contenu dans votre source de données
• maxFileSizeInMegaBytes	Spécifiez une limite de taille MBs de Amazon Kendra fichier à analyser.
• Commentaire iScrawl	`true`pour analyser les commentaires dans votre source de données Google Drive.
• isCrawlMyDriveAndSharedWithMe	`true`pour explorer MyDrive et partager des disques avec moi dans votre source de données Google Drive.
• isCrawlSharedDisques	`true`pour explorer les lecteurs partagés dans votre source de données Google Drive.

Configuration	Description
isCrawlAcl	`true`pour analyser les informations de la liste de contrôle d'accès (ACL) de vos documents , si vous disposez d'une ACL et que vous souhaitez l'utiliser pour le contrôle d'accès. L'ACL indique les documents auxquels les utilisateurs et les groupes peuvent accéder et effectuer des recherches. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.
• excludeUserAccounts • excludeSharedDrives • excludeMimeTypes • exclusionFileTypeMotifs • exclusionFileNameMotifs • exclusionFilePathFiltre	Liste de modèles d'expressions régulières permettant d'exclure certains fichiers de votre source de données Google Drive. Les fichiers qui correspondent aux modèles sont exclus de l'index. Les fichiers qui ne correspondent pas aux modèles sont inclus dans l'index. Si un fichier correspond à la fois à un modèle d'exclusion et à un modèle d'inclusion, le modèle d'exclusion est prioritaire et le fichier n'est pas inclus dans l'index.
• includeUserAccounts • includeSharedDrives • includeMimeTypes • inclusionFileTypeMotifs • inclusionFileNameMotifs • inclusionFilePathFiltre	Liste de modèles d'expressions régulières permettant d'inclure certains fichiers dans votre source de données Google Drive. Les fichiers qui correspondent aux modèles sont inclus dans l'index. Les fichiers qui ne correspon dent pas aux modèles sont exclus de l'index. Si un fichier correspond à la fois à un modèle d'inclusion et à un modèle d'exclusion, le modèle d'exclusion est prioritaire et le fichier n'est pas inclus dans l'index.

Configuration	Description
type	Type de source de données. Spécifiez `GOOGLEDRIVEV2` comme type de source de données.
enableIdentityCrawler	`true`utiliser Amazon Kendra le moteur de recherche d'identité pour synchroniser les identity/principal informations relatives aux utilisateurs et aux groupes ayant accès à certains documents. Si le robot d'identification est désactivé, tous les documents peuvent être consultés publiquement. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'identification est désactivé , vous pouvez également utiliser l'PutPrincipalMappingAPI pour télécharger les informations d'accès des utilisateurs et des groupes.

Configuration	Description
Mode de synchronisation	Spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Vous pouvez choisir entre : • `FORCED_FULL_CRAWL` pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index. • `FULL_CRAWL` pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modificat ions de contenu et indexer le contenu modifié depuis la dernière synchronisation. • `CHANGE_LOG` pour indexer uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchroni se avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

Configuration	Description

Secrétaire Arn	Le nom de ressource Amazon (ARN) d'un AWS Secrets Manager secret qui contient les paires clé-valeur requises pour se connecter à votre Google Drive. Le secret doit contenir une structure JSON avec les clés suivantes :
	
	Si vous utilisez l'authentification par compte de service Google :
	
	`{ "clientEmail": " user account email", "adminAccountEmail": " service account email", "privateKey": " private key" }`
	
	Si vous utilisez l'authentification OAuth 2.0 :
	
	`{ "clientID": " OAuth client ID", "clientSecret": " client secret", "refreshToken": " refresh token" }`

version	Version de ce modèle actuellement prise en charge.

Schéma JSON de Google Drive

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
```

```
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "authType": {
              "type": "string",
              "enum": [
                "serviceAccount",
                "OAuth2"
              ]
            }
          },
          "required": [
            "authType"
          ]
        }
      },
      "required": [
        "repositoryEndpointMetadata"
      ]
    },
    "repositoryConfigurations": {
      "type": "object",
      "properties": {
        "file": {
          "type": "object",
          "properties": {
            "fieldMappings": {
              "type": "array",
              "items": [
                {
                  "type": "object",
                  "properties": {
                    "indexFieldName": {
                      "type": "string"
                    },
                    "indexFieldType": {
                      "type": "string",
                      "enum": [
                        "STRING",
                        "DATE",
                        "STRING_LIST",
                        "LONG"
                      ]
                    },
```

```
                "dataSourceFieldName": {
                  "type": "string"
                },
                "dateFieldFormat": {
                  "type": "string",
                  "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                }
              },
              "required": [
                "indexFieldName",
                "indexFieldType",
                "dataSourceFieldName"
              ]
            }
          ]
        }
      },
      "required": [
        "fieldMappings"
      ]
    },
    "comment": {
      "type": "object",
      "properties": {
        "fieldMappings": {
          "type": "array",
          "items": [
            {
              "type": "object",
              "properties": {
                "indexFieldName": {
                  "type": "string"
                },
                "indexFieldType": {
                  "type": "string",
                  "enum": [
                    "STRING",
                    "DATE",
                    "STRING_LIST"
                  ]
                },
                "dataSourceFieldName": {
                  "type": "string"
                },
```

```
                  "dateFieldFormat": {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required": [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      }
    }
  },
  "additionalProperties": {
    "type": "object",
    "properties": {
      "maxFileSizeInMegaBytes": {
        "type": "string"
      },
      "isCrawlComment": {
        "type": "boolean"
      },
      "isCrawlMyDriveAndSharedWithMe": {
        "type": "boolean"
      },
      "isCrawlSharedDrives": {
        "type": "boolean"
      },
      "isCrawlAcl": {
        "type": "boolean"
      },
      "excludeUserAccounts": {
        "type": "array",
        "items": {
          "type": "string"
        }
      },
```

```
        "excludeSharedDrives": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "excludeMimeTypes": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "includeUserAccounts": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "includeSharedDrives": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "includeMimeTypes": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "includeTargetAudienceGroup": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "inclusionFileTypePatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "inclusionFileNamePatterns": {
          "type": "array",
```

```
          "items": {
            "type": "string"
          }
        },
        "exclusionFileTypePatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "exclusionFileNamePatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "inclusionFilePathFilter": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "exclusionFilePathFilter": {
          "type": "array",
          "items": {
            "type": "string"
          }
        }
      }
    },
    "type": {
      "type": "string",
      "pattern": "GOOGLEDRIVEV2"
    },
    "enableIdentityCrawler": {
      "type": "boolean"
    },
    "syncMode": {
      "type": "string",
      "enum": [
        "FORCED_FULL_CRAWL",
        "FULL_CRAWL",
        "CHANGE_LOG"
      ]
```

```
    },
    "secretArn": {
      "type": "string",
      "minLength": 20,
      "maxLength": 2048
    }
  },
  "version": {
    "type": "string",
    "anyOf": [
      {
        "pattern": "1.0.0"
      }
    ]
  },
  "required": [
    "connectionConfiguration",
    "repositoryConfigurations",
    "syncMode",
    "additionalProperties",
    "secretArn",
    "type"
  ]
}
```

Schéma DB2 du modèle IBM

Vous incluez un JSON qui contient le schéma de la source de données dans le cadre de l'TemplateConfigurationobjet. Spécifiez le type de source de donnéesJDBC, le type de base de donnéesdb2, un secret pour vos informations d'authentification et les autres configurations nécessaires. Vous spécifiez ensuite TEMPLATE le Type moment de l'appel CreateDataSource.

Vous pouvez utiliser le modèle fourni dans ce guide du développeur. Consultez Schéma IBM DB2 JSON.

Le tableau suivant décrit les paramètres du schéma IBM DB2 JSON.

Configuration	Description
Configuration de connexion	Informations de configuration pour le point de terminaison de la source de données.

Configuration	Description
repositoryEndpointMetadata	Informations de configuration requises pour connecter votre source de données. • DBType : le type de base de données Java que vous utilisez, que ce soit, `mysql` `db2`, `postgresql` ou. `oracle` `sqlserver` • DBHost : nom d'hôte de la base de données. • DBPort : port de base de données. • DBInstance : instance de base de données.
Configurations du référentiel	Informations de configuration pour le contenu de la source de données. Par exemple, configurer des types spécifiques de contenu et des mappages de champs. Spécifiez le type de source de données et l'ARN secret.
document	Liste d'objets qui mappent les attributs ou les noms de champs du contenu de votre base de données aux noms de champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.
Propriétés supplémentaires	Options de configuration supplémentaires pour votre contenu dans votre source de données. À utiliser pour inclure ou exclure un contenu spécifique dans la source de données de votre base de données.
Clé primaire	Fournissez la clé primaire pour la table de base de données. Cela permet d'identifier une table au sein de votre base de données.

Configuration	Description
Colonne de titre	Indiquez le nom de la colonne du titre du document dans votre table de base de données.
Colonne Body	Indiquez le nom de la colonne du titre du document dans votre table de base de données.
sqlQuery	Entrez des instructions de requête SQL telles que les opérations SELECT et JOIN. La taille des requêtes SQL doit être inférieure à 32 Ko. Amazon Kendra analysera tout le contenu de la base de données correspondant à votre requête.
Colonne d'horodatage	Entrez le nom de la colonne contenant les horodatages. Amazon Kendra utilise les informations d'horodatage pour détecter les modifications apportées à votre contenu et synchroniser uniquement le contenu modifié.
Format d'horodatage	Entrez le nom de la colonne contenant les formats d'horodatage à utiliser pour détecter les modifications de contenu et resynchroniser votre contenu.
timezone	Entrez le nom de la colonne contenant les fuseaux horaires du contenu à analyser.
changeDetectingColumns	Entrez le nom des colonnes qui Amazon Kendra seront utilisées pour détecter les modifications de contenu. Amazon Kendra réindexera le contenu en cas de modification de l'une de ces colonnes

Configuration	Description
allowedUsersColumns	Entrez le nom de la colonne contenant l'utilisa teur devant IDs être autorisé à accéder au contenu.
allowedGroupsColumn	Entrez le nom de la colonne contenant l'utilisa teur devant IDs être autorisé à accéder au contenu.
source URIColumn	Entrez le nom de la colonne contenant la source URLs à indexer.
isSslEnabled	Entrez des instructions de requête SQL telles que les opérations SELECT et JOIN. La taille des requêtes SQL doit être inférieure à 32 Ko. Amazon Kendra analysera tout le contenu de la base de données correspondant à votre requête.
type	Type de source de données. Spécifiez `JDBC` comme type de source de données.

Configuration	Description
Mode de synchronisation	Spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Vous pouvez choisir entre : • `FORCED_FULL_CRAWL` pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index. • `FULL_CRAWL` pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modificat ions de contenu et indexer le contenu modifié depuis la dernière synchronisation. • `CHANGE_LOG` pour indexer uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchroni se avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

Configuration	Description
Secrétaire Arn	Le nom de ressource Amazon (ARN) d'un secret Secrets Manager contenant le nom d'utilisateur et le mot de passe requis pour se connecter à votre base de données. Le secret doit contenir une structure JSON avec les clés suivantes : `{ "user name": "database user name", "password": " password" }`
version	Version du modèle actuellement prise en charge.

Schéma IBM DB2 JSON

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "dbType": {
              "type": "string",
              "enum": [
                "mysql",
                "db2",
                "postgresql",
                "oracle",
                "sqlserver"
              ]
            },
            "dbHost": {
```

```
            "type": "string"
          },
          "dbPort": {
            "type": "string"
          },
          "dbInstance": {
            "type": "string"
          }
        },
        "required": [
          "dbType",
          "dbHost",
          "dbPort",
          "dbInstance"
        ]
      }
    },
    "required": [
      "repositoryEndpointMetadata"
    ]
  },
  "repositoryConfigurations": {
    "type": "object",
    "properties": {
      "document": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string"
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  }
                },
                "required": [
```

```
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      }
    },
    "required": [
    ]
  },
  "additionalProperties": {
    "type": "object",
    "properties": {
      "primaryKey": {
        "type": "string"
      },
      "titleColumn": {
        "type": "string"
      },
      "bodyColumn": {
        "type": "string"
      },
      "sqlQuery": {
        "type": "string",
        "not": {
          "pattern": ";+"
        }
      },
      "timestampColumn": {
        "type": "string"
      },
      "timestampFormat": {
        "type": "string"
      },
      "timezone": {
        "type": "string"
      },
      "changeDetectingColumns": {
```

```
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "allowedUsersColumn": {
          "type": "string"
        },
        "allowedGroupsColumn": {
          "type": "string"
        },
        "sourceURIColumn": {
          "type": "string"
        },
        "isSslEnabled": {
          "type": "boolean"
        }
      },
      "required": ["primaryKey", "titleColumn", "bodyColumn", "sqlQuery"]
    },
    "type" : {
      "type" : "string",
      "pattern": "JDBC"
    },
    "syncMode": {
      "type": "string",
      "enum": [
        "FORCED_FULL_CRAWL",
        "FULL_CRAWL",
        "CHANGE_LOG"
      ]
    },
    "secretArn": {
      "type": "string"
    }
  },
  "version": {
    "type": "string",
    "anyOf": [
      {
        "pattern": "1.0.0"
      }
    ]
  },
```

```
  "required": [
      "connectionConfiguration",
      "repositoryConfigurations",
      "syncMode",
      "additionalProperties",
      "secretArn",
      "type"
  ]
}
```

Schéma du modèle Microsoft Exchange

Vous incluez un JSON qui contient le schéma de la source de données dans le cadre de l'TemplateConfigurationobjet. Vous fournissez l'ID du locataire dans le cadre de la configuration de la connexion ou des détails du point de terminaison du référentiel. Spécifiez également le type de source de donnéesMSEXCHANGE, un secret pour vos informations d'authentification et les autres configurations nécessaires. Vous spécifiez ensuite TEMPLATE le Type moment de l'appel CreateDataSource.

Vous pouvez utiliser le modèle fourni dans ce guide du développeur. Consultez Schéma Microsoft Exchange JSON.

Le tableau suivant décrit les paramètres du schéma Microsoft Exchange JSON.

Configuration	Description
Configuration de connexion	Informations de configuration pour le point de terminaison de la source de données.
repositoryEndpointMetadata	Informations sur le point de terminaison de la source de données.
ID du locataire	L'identifiant du client Microsoft 365. Vous pouvez trouver votre ID de locataire dans les propriétés de votre portail Azure Active Directory ou dans votre OAuth application.
Configurations du référentiel	Informations de configuration pour le contenu de la source de données. Par exemple,

Configuration	Description
	configurer des types spécifiques de contenu et des mappages de champs.
• e-mail • attachment • calendrier • contacts • notes	Liste d'objets qui mappent les attributs ou les noms de champs de votre source de données Microsoft Exchange à des champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.
Propriétés supplémentaires	Options de configuration supplémentaires pour le contenu de votre source de données
Modèles d'inclusion	Liste de modèles d'expressions régulières permettant d'inclure certains fichiers dans votre source de données Microsoft Exchange. Les fichiers qui correspondent aux modèles sont inclus dans l'index. Les fichiers qui ne correspondent pas aux modèles sont exclus de l'index. Si un fichier correspond à la fois à un modèle d'inclusion et d'exclusion, le modèle d'exclusion a la priorité et le fichier n'est pas inclus dans l'index.
Schémas d'exclusion	Liste de modèles d'expressions régulières permettant d'exclure certains fichiers de votre source de données Microsoft Exchange. Les fichiers qui correspondent aux modèles sont exclus de l'index. Les fichiers qui ne correspon dent pas aux modèles sont inclus dans l'index. Si un fichier correspond à la fois à un modèle d'exclusion et à un modèle d'inclusion, le modèle d'exclusion est prioritaire et le fichier n'est pas inclus dans l'index.

Configuration	Description
• inclusionUsersList • inclusionUsersFileNom • inclusionDomainUsers	Liste de modèles d'expressions régulière s permettant d'inclure certains utilisateurs et fichiers utilisateur dans votre source de données Microsoft Exchange. Les utilisateurs qui correspondent aux modèles sont inclus dans l'index. Les utilisateurs qui ne correspon dent pas aux modèles sont exclus de l'index. Si un utilisateur correspond à la fois à un modèle d'inclusion et à un modèle d'exclusion, le modèle d'exclusion est prioritaire et l'utilisateur n'est pas inclus dans l'index.
• exclusionUsersList • exclusionUsersFileNom • exclusionDomainUsers	Liste de modèles d'expressions régulières permettant d'exclure certains utilisateurs et fichiers utilisateur de votre source de données Microsoft Exchange. Les utilisateurs qui correspondent aux modèles sont exclus de l'index. Les utilisateurs qui ne correspondent pas aux modèles sont inclus dans l'index. Si un utilisateur correspond à la fois à un modèle d'exclusion et à un modèle d'inclusion, le modèle d'exclusion est prioritaire et l'utilisateur n'est pas inclus dans l'index.
Nom du compartiment S3	Le nom de votre compartiment S3 si vous souhaitez l'utiliser.
• Calendrier Crawl • Notes d'exploration • Analyser les contacts • crawlFolderAcl	`true`pour analyser ces types de contenu et d'informations de contrôle d'accès dans votre source de données Microsoft Exchange.
startCalendarDateHeure	Vous pouvez configurer une date et une heure de début spécifiques pour le contenu de votre calendrier.

Configuration	Description
endCalendarDateHeure	Vous pouvez configurer une date et une heure de fin spécifiques pour le contenu du calendrier.
subject	Vous pouvez configurer une ligne d'objet spécifique pour le contenu de votre e-mail.
Courrier électronique de	Vous pouvez configurer un e-mail spécifique pour le contenu de votre « expéditeur » ou du courrier de l'expéditeur.
Envoyer un e-mail à	Vous pouvez configurer un e-mail spécifique pour le contenu de votre message « À » ou du message du destinataire.

Configuration	Description
Mode de synchronisation	Spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Vous pouvez choisir entre : • `FORCED_FULL_CRAWL` pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index. • `FULL_CRAWL` pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modificat ions de contenu et indexer le contenu modifié depuis la dernière synchronisation. • `CHANGE_LOG` pour indexer uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchroni se avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
type	Type de source de données. Spécifiez `MSEXCHANGE` comme type de source de données.

Configuration	Description
Secrétaire Arn	Le nom de ressource Amazon (ARN) d'un AWS Secrets Manager secret qui contient les paires clé-valeur requises pour se connecter à votre Microsoft Exchange. Cela inclut votre ID client et votre secret client générés lorsque vous créez une OAuth application sur le portail Azure.
version	Version de ce modèle actuellement prise en charge.

Schéma Microsoft Exchange JSON

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "tenantId": {
              "type": "string",
              "pattern": "^[0-9a-f]{8}-[0-9a-f]{4}-[0-9a-f]{4}-[0-9a-f]{4}-[0-9a-f]
{12}$",
              "minLength": 36,
              "maxLength": 36
            }
          },
          "required": ["tenantId"]
        }
      }
    },
    "repositoryConfigurations": {
      "type": "object",
      "properties": {
```

```
        "email": {
          "type": "object",
          "properties": {
            "fieldMappings": {
              "type": "array",
              "items": [
                {
                  "type": "object",
                  "properties": {
                    "indexFieldName": {
                      "type": "string"
                    },
                    "indexFieldType": {
                      "type": "string",
                      "enum": ["STRING", "STRING_LIST", "DATE"]
                    },
                    "dataSourceFieldName": {
                      "type": "string"
                    },
                    "dateFieldFormat": {
                      "type": "string",
                      "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                    }
                  },
                  "required": [
                    "indexFieldName",
                    "indexFieldType",
                    "dataSourceFieldName"
                  ]
                }
              ]
            }
          },
          "required": [
            "fieldMappings"
          ]
        },
        "attachment": {
          "type": "object",
          "properties": {
            "fieldMappings": {
              "type": "array",
              "items": [
                {
```

```
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string",
                    "enum": ["STRING", "DATE","LONG"]
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  },
                  "dateFieldFormat": {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required": [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      },
      "calendar": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string",
```

```
                    "enum": ["STRING", "STRING_LIST", "DATE"]
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  },
                  "dateFieldFormat": {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required": [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      },
      "contacts": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string",
                    "enum": ["STRING", "STRING_LIST", "DATE"]
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  },
                  "dateFieldFormat": {
                    "type": "string",
```

```
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required": [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      },
      "notes": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string",
                    "enum": ["STRING", "DATE"]
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  },
                  "dateFieldFormat": {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required": [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
```

```
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      }
    },
    "required": ["email"
    ]
  },
  "additionalProperties": {
    "type": "object",
    "properties": {
      "inclusionPatterns": {
        "type": "array",
        "items": {
          "type": "string"
        }
      },
      "exclusionPatterns": {
        "type": "array",
        "items": {
          "type": "string"
        }
      },
      "inclusionUsersList": {
        "type": "array",
        "items": {
          "type": "string",
          "format": "email"
        }
      },
      "exclusionUsersList": {
        "type": "array",
        "items": {
          "type": "string",
          "format": "email"
        }
      },
      "s3bucketName": {
        "type": "string"
```

```
        },
        "inclusionUsersFileName": {
          "type": "string"
        },
        "exclusionUsersFileName": {
          "type": "string"
        },
        "inclusionDomainUsers": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "exclusionDomainUsers": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "crawlCalendar": {
          "type": "boolean"
        },
        "crawlNotes": {
          "type": "boolean"
        },
        "crawlContacts": {
          "type": "boolean"
        },
        "crawlFolderAcl": {
          "type": "boolean"
        },
        "startCalendarDateTime": {
          "anyOf": [
            {
              "type": "string",
              "pattern": "^[0-9]{4}-[0-9]{2}-[0-9]{2}T[0-9]{2}:[0-9]{2}:[0-9]{2}Z$"
            },
            {
              "type": "string",
              "pattern": ""
            }
          ]
        },
        "endCalendarDateTime": {
```

```
          "anyOf": [
            {
            "type": "string",
            "pattern": "^[0-9]{4}-[0-9]{2}-[0-9]{2}T[0-9]{2}:[0-9]{2}:[0-9]{2}Z$"
            },
            {
              "type": "string",
              "pattern": ""
            }
          ]
        },
        "subject": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "emailFrom": {
          "type": "array",
          "items": {
            "type": "string",
            "format": "email"
          }
        },
        "emailTo": {
          "type": "array",
          "items": {
            "type": "string",
            "format": "email"
          }
        }
      },
      "required": [
      ]
    },
    "syncMode": {
      "type": "string",
      "enum": [
        "FORCED_FULL_CRAWL",
        "FULL_CRAWL",
        "CHANGE_LOG"
      ]
    },
    "type" : {
```

```
      "type" : "string",
      "pattern": "MSEXCHANGE"
    },
    "secretArn": {
      "type": "string"
    }
  },
  "version": {
    "type": "string",
    "anyOf": [
      {
        "pattern": "1.0.0"
      }
    ]
  },
  "required": [
    "connectionConfiguration",
    "repositoryConfigurations",
    "syncMode",
    "additionalProperties",
    "secretArn",
    "type"
  ]
}
```

Schéma OneDrive de modèle Microsoft

Vous incluez un JSON qui contient le schéma de la source de données dans le cadre de l'TemplateConfigurationobjet. Vous fournissez l'ID du locataire dans le cadre de la configuration de la connexion ou des détails du point de terminaison du référentiel. Spécifiez également le type de source de données et un secret pour vos informations d'authentification, ainsi que les autres configurations nécessaires. `ONEDRIVEV2` Vous spécifiez ensuite `TEMPLATE` le `Type` moment de l'appel CreateDataSource.

Vous pouvez utiliser le modèle fourni dans ce guide du développeur. Consultez Schéma Microsoft OneDrive JSON.

Le tableau suivant décrit les paramètres du schéma Microsoft OneDrive JSON.

Configuration	Description
Configuration de connexion	Informations de configuration pour le point de terminaison de la source de données.
repositoryEndpointMetadata	Informations sur le point de terminaison de la source de données.
ID du locataire	L'identifiant du client Microsoft 365. Vous pouvez trouver votre ID de locataire dans les propriétés de votre portail Azure Active Directory ou dans votre OAuth application.
Configurations du référentiel	Informations de configuration pour le contenu de la source de données. Par exemple, configurer des types spécifiques de contenu et des mappages de champs.
dans le fichier	Liste d'objets qui mappent les attributs ou les noms de champs de vos OneDrive fichiers Microsoft aux noms de champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.
Propriétés supplémentaires	Options de configuration supplémentaires pour votre contenu dans votre source de données
• userNameFilter • userFilterPath • inclusionFileTypeMotifs • exclusionFileTypeMotifs • inclusionFileNameMotifs • exclusionFileNameMotifs • inclusionFilePathMotifs • exclusionFilePathMotifs	Vous pouvez choisir d'indexer des fichiers, des OneNote sections, des OneNote pages spécifiques et de filtrer par nom d'utilisateur.

Configuration	Description
• inclusionOneNoteSectionNamePatterns • exclusionOneNoteSectionNamePatterns • inclusionOneNotePageNamePatterns • exclusionOneNotepageNamePatterns	
isUserNamesur S3	`true`pour fournir une liste de noms d'utilisateur dans un fichier stocké dans un fichier Amazon S3.
type	Type de source de données. Spécifiez `ONEDRIVEV2` comme type de source de données.
enableIdentityCrawler	`true`utiliser Amazon Kendra le moteur de recherche d'identité pour synchroniser les identity/principal informations relatives aux utilisateurs et aux groupes ayant accès à certains documents. Si le robot d'identification est désactivé, tous les documents peuvent être consultés publiquement. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'identification est désactivé , vous pouvez également utiliser l'PutPrincipalMappingAPI pour télécharger les informations d'accès des utilisateurs et des groupes.
type	Type de source de données. Spécifiez `ONEDRIVEV2` comme type de source de données.

Configuration	Description
Mode de synchronisation	Spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Vous pouvez choisir entre : • `FORCED_FULL_CRAWL` pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index. • `FULL_CRAWL` pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation. • `CHANGE_LOG` pour indexer uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

Configuration	Description
Secrétaire Arn	Le nom de ressource Amazon (ARN) d'un AWS Secrets Manager secret qui contient les paires clé-valeur requises pour se connecter à votre Microsoft. OneDrive Le secret doit contenir une structure JSON avec les clés suivantes : `{ "clientId": " client ID", "clientSecret": " client secret" }`
version	Version de ce modèle actuellement prise en charge.

Schéma Microsoft OneDrive JSON

```
{
 "$schema": "http://json-schema.org/draft-04/schema#",
 "type": "object",
 "properties": {
  "connectionConfiguration": {
   "type": "object",
   "properties": {
    "repositoryEndpointMetadata": {
     "type": "object",
     "properties": {
      "tenantId": {
       "type": "string",
       "pattern": "^[0-9a-f]{8}-[0-9a-f]{4}-[0-9a-f]{4}-[0-9a-f]{4}-[0-9a-f]{12}$",
       "minLength": 36,
       "maxLength": 36
      }
     },
     "required": [
      "tenantId"
     ]
    }
   },
```

```
   "required": [
    "repositoryEndpointMetadata"
   ]
  },
  "repositoryConfigurations": {
   "type": "object",
   "properties": {
    "file": {
     "type": "object",
     "properties": {
      "fieldMappings": {
       "type": "array",
       "items": [
        {
         "type": "object",
         "properties": {
          "indexFieldName": {
           "type": "string"
          },
          "indexFieldType": {
           "type": "string",
           "enum": [
            "STRING",
            "STRING_LIST",
            "DATE",
            "LONG"
           ]
          },
          "dataSourceFieldName": {
           "type": "string"
          },
          "dateFieldFormat": {
           "type": "string",
           "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
          }
         },
         "required": [
          "indexFieldName",
          "indexFieldType",
          "dataSourceFieldName"
         ]
        }
       ]
      }
```

```
     },
     "required": [
      "fieldMappings"
     ]
    }
   }
  },
  "additionalProperties": {
   "type": "object",
   "properties": {
    "userNameFilter": {
     "type": "array",
     "items": {
      "type": "string"
     }
    },
    "userFilterPath": {
     "type": "string"
    },
    "isUserNameOnS3": {
     "type": "boolean"
    },
    "inclusionFileTypePatterns": {
     "type": "array",
     "items": {
      "type": "string"
     }
    },
    "exclusionFileTypePatterns": {
     "type": "array",
     "items": {
      "type": "string"
     }
    },
    "inclusionFileNamePatterns": {
     "type": "array",
     "items": {
      "type": "string"
     }
    },
    "exclusionFileNamePatterns": {
     "type": "array",
     "items": {
      "type": "string"
```

```
     }
    },
    "inclusionFilePathPatterns": {
     "type": "array",
     "items": {
      "type": "string"
     }
    },
    "exclusionFilePathPatterns": {
     "type": "array",
     "items": {
      "type": "string"
     }
    },
    "inclusionOneNoteSectionNamePatterns": {
     "type": "array",
     "items": {
      "type": "string"
     }
    },
    "exclusionOneNoteSectionNamePatterns": {
     "type": "array",
     "items": {
      "type": "string"
     }
    },
    "inclusionOneNotePageNamePatterns": {
     "type": "array",
     "items": {
      "type": "string"
     }
    },
    "exclusionOneNotePageNamePatterns": {
     "type": "array",
     "items": {
      "type": "string"
     }
    }
   },
   "required": []
  },

  "enableIdentityCrawler": {
   "type": "boolean"
```

```
  },
  "type": {
   "type": "string",
   "pattern": "ONEDRIVEV2"
  },
  "syncMode": {
   "type": "string",
   "enum": [
    "FULL_CRAWL",
    "FORCED_FULL_CRAWL",
    "CHANGE_LOG"
   ]
  },
  "secretArn": {
   "type": "string",
   "minLength": 20,
   "maxLength": 2048
  }
 },
 "version": {
  "type": "string",
  "anyOf": [
   {
    "pattern": "1.0.0"
   }
  ]
 },
 "required": [
  "connectionConfiguration",
  "repositoryConfigurations",
  "syncMode",
  "additionalProperties",
  "secretArn",
  "type"
 ]
}
```

Schéma SharePoint de modèle Microsoft

Vous incluez un JSON qui contient le schéma de la source de données dans le cadre de TemplateConfigurationl'objet. Vous fournissez l'URL du SharePoint siteURLs, le domaine et également un ID de locataire si nécessaire dans le cadre de la configuration de la connexion ou des détails du point de terminaison du référentiel. Spécifiez également le type de source

de donnéesSHAREPOINTV2, un secret pour vos informations d'authentification et les autres configurations nécessaires. Vous le spécifiez ensuite TEMPLATE comme type lorsque vous appelez CreateDataSource.

Vous pouvez utiliser le modèle fourni dans ce guide du développeur. Consultez SharePoint Schéma JSON.

Le tableau suivant décrit les paramètres du schéma Microsoft SharePoint JSON.

Configuration	Description
Configuration de connexion	Informations de configuration du point de terminaison pour la source de données
repositoryEndpointMetadata	Informations sur le point de terminaison pour la source de données
ID du locataire	L'identifiant du locataire de votre SharePoint compte.
domaine	Le domaine de votre SharePoint compte.
URL du site	L'hébergeur URLs de votre SharePoint compte.
repositoryAdditionalProperties	Propriétés supplémentaires pour se connecter au point de terminaison repository/data source.
Nom du compartiment S3	Nom du Amazon S3 compartiment qui stocke votre certificat X.509 autosigné Azure AD.
Nom du certificat S3	Le nom du certificat X.509 autosigné Azure AD stocké dans votre Amazon S3 compartiment.
authType	Le type d'authentification que vous utilisez, qu'il s'agisse de OAuth2 OAuth2Certificate OAuth2App ,Basic,OAuth2_RefreshToken ,NTLM, ouKerberos.
version	La SharePoint version que vous utilisez, que ce soit Server ouOnline.

Configuration	Description
onPremVersion	La version SharePoint du serveur que vous utilisez, que ce soit `2013` `20162019`, ou`SubscriptionEdition` .
Configurations du référentiel	Informations de configuration pour le contenu de la source de données. Par exemple, configurer des types spécifiques de contenu et des mappages de champs.
• event • page • dans le fichier • lien • attachment • comment	Liste d'objets qui mappent les attributs ou les noms de champs de votre SharePoint contenu aux noms de champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.
Propriétés supplémentaires	Options de configuration supplémentaires pour votre contenu dans votre source de données.
• eventTitleFilterRegEx • pageTitleFilterRegEx • linkTitleFilterRegEx • inclusionFilePath • exclusionFilePath • inclusionFileTypeMotifs • exclusionFileTypeMotifs • inclusionFileNameMotifs • exclusionFileNameMotifs • inclusionOneNoteSectionNamePatterns • exclusionOneNoteSectionNamePatterns • inclusionOneNotePageNamePatterns • exclusionOneNotePageNamePatterns	Liste de modèles d'expressions régulières pour include/exclude certains contenus de votre source de SharePoint données. Les éléments de contenu qui correspondent aux modèles d'inclusion sont inclus dans l'index. Les éléments de contenu qui ne correspondent pas aux modèles d'inclusion sont exclus de l'index. Si un fichier correspond à la fois à un modèle d'inclusion et à un modèle d'exclusion, le modèle d'exclusion est prioritaire et le fichier n'est pas inclus dans l'index.

Configuration	Description
• Fichiers d'exploration • Parcourir les pages • Évènements sur chenilles • Explorez les commentaires • Liens d'exploration • Accessoires Crawl	`true`pour explorer ces types de contenu.
CrawlacL	`true`pour analyser les informations de la liste de contrôle d'accès (ACL) de vos documents , si vous disposez d'une ACL et que vous souhaitez l'utiliser pour le contrôle d'accès. L'ACL indique les documents auxquels les utilisateurs et les groupes peuvent accéder et effectuer des recherches. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.
fieldForUserID	Spécifiez `email` si vous souhaitez utiliser l'adresse e-mail de l'utilisateur pour l'ID utilisateur ou `userPrincipalName` si vous souhaitez utiliser un nom d'utilisateur pour l'ID utilisateur. Si vous ne spécifiez aucune option, elle `email` est utilisée par défaut.
Configuration de l'ACL	Spécifiez soit `ACLWithLDAPEmailFmt` `ACLWithManualEmailFmt` , soit`ACLWithUsernameFmtM` .
Domaine de messagerie	Le domaine de l'e-mail. Par exemple, « *`amazon.com`* ».

Configuration	Description
• isCrawlLocalGroupMapping • isCrawlAdGroupMapping	`true`pour analyser les informations de mappage des groupes.
Hôte proxy	Le nom d'hôte du proxy Web que vous utilisez, sans le protocole http ://ou https ://.
Port proxy	Numéro de port utilisé par le protocole de transport d'URL de l'hôte. Il doit s'agir d'une valeur numérique comprise entre 0 et 65535.
type	Spécifiez `SHAREPOINTV2` comme type de source de données
enableIdentityCrawler	`true`utiliser Amazon Kendra le moteur de recherche d'identité pour synchroniser les identity/principal informations relatives aux utilisateurs et aux groupes ayant accès à certains documents. Si le robot d'identification est désactivé, tous les documents peuvent être consultés publiquement. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'identification est désactivé , vous pouvez également utiliser l'PutPrincipalMappingAPI pour télécharger les informations d'accès des utilisateurs et des groupes.

Configuration	Description
Mode de synchronisation	Spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Vous pouvez choisir entre : • `FORCED_FULL_CRAWL` pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index. • `FULL_CRAWL` pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modificat ions de contenu et indexer le contenu modifié depuis la dernière synchronisation. • `CHANGE_LOG` pour indexer uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchroni se avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
Secrétaire Arn	Le nom de ressource Amazon (ARN) d'un AWS Secrets Manager secret contenant les paires clé-valeur requises pour se connecter à votre. SharePoint Pour plus d'informations sur ces paires clé-valeur, consultez les instructi ons de connexion pour SharePoint Online et SharePoint Server.

Configuration	Description
version	Version de ce modèle actuellement prise en charge.

SharePoint Schéma JSON

```
{
 "$schema": "http://json-schema.org/draft-04/schema#",
 "type": "object",
 "properties": {
  "connectionConfiguration": {
   "type": "object",
   "properties": {
    "repositoryEndpointMetadata": {
     "type": "object",
     "properties": {
      "tenantId": {
       "type": "string",
       "pattern": "^[0-9a-f]{8}-[0-9a-f]{4}-[0-9a-f]{4}-[0-9a-f]{4}-[0-9a-f]{12}$",
       "minLength": 36,
       "maxLength": 36
      },
      "domain": {
       "type": "string"
      },
      "siteUrls": {
       "type": "array",
       "items": {
        "type": "string",
        "pattern": "https://.*"
       }
      },
      "repositoryAdditionalProperties": {
       "type": "object",
       "properties": {
        "s3bucketName": {
         "type": "string"
        },
        "s3certificateName": {
         "type": "string"
        },
```

```
        "authType": {
         "type": "string",
         "enum": [
          "OAuth2",
          "OAuth2Certificate",
          "OAuth2App",
          "Basic",
          "OAuth2_RefreshToken",
          "NTLM",
          "Kerberos"
         ]
        },
        "version": {
         "type": "string",
         "enum": [
          "Server",
          "Online"
         ]
        },
        "onPremVersion": {
         "type": "string",
         "enum": [
          "",
          "2013",
          "2016",
          "2019",
          "SubscriptionEdition"
         ]
        }
       },
       "required": [
        "authType",
        "version"
       ]
      }
     },
     "required": [
      "siteUrls",
      "domain",
      "repositoryAdditionalProperties"
     ]
    }
   },
   "required": [
```

```
    "repositoryEndpointMetadata"
   ]
  },
  "repositoryConfigurations": {
   "type": "object",
   "properties": {
    "event": {
     "type": "object",
     "properties": {
      "fieldMappings": {
       "type": "array",
       "items": [
        {
         "type": "object",
         "properties": {
          "indexFieldName": {
           "type": "string"
          },
          "indexFieldType": {
           "type": "string",
           "enum": [
            "STRING",
            "STRING_LIST",
            "DATE"
           ]
          },
          "dataSourceFieldName": {
           "type": "string"
          },
          "dateFieldFormat": {
           "type": "string",
           "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
          }
         },
         "required": [
          "indexFieldName",
          "indexFieldType",
          "dataSourceFieldName"
         ]
        }
       ]
      }
     },
     "required": [
```

```
      "fieldMappings"
     ]
    },
    "page": {
     "type": "object",
     "properties": {
      "fieldMappings": {
       "type": "array",
       "items": [
        {
         "type": "object",
         "properties": {
          "indexFieldName": {
           "type": "string"
          },
          "indexFieldType": {
           "type": "string",
           "enum": [
            "STRING",
            "DATE",
            "LONG"
           ]
          },
          "dataSourceFieldName": {
           "type": "string"
          },
          "dateFieldFormat": {
           "type": "string",
           "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
          }
         },
         "required": [
          "indexFieldName",
          "indexFieldType",
          "dataSourceFieldName"
         ]
        }
       ]
      }
     },
     "required": [
      "fieldMappings"
     ]
    },
```

```
    "file": {
     "type": "object",
     "properties": {
      "fieldMappings": {
       "type": "array",
       "items": [
        {
         "type": "object",
         "properties": {
          "indexFieldName": {
           "type": "string"
          },
          "indexFieldType": {
           "type": "string",
           "enum": [
            "STRING",
            "DATE",
            "LONG"
           ]
          },
          "dataSourceFieldName": {
           "type": "string"
          },
          "dateFieldFormat": {
           "type": "string",
           "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
          }
         },
         "required": [
          "indexFieldName",
          "indexFieldType",
          "dataSourceFieldName"
         ]
        }
       ]
      }
     },
     "required": [
      "fieldMappings"
     ]
    },
    "link": {
     "type": "object",
     "properties": {
```

```
      "fieldMappings": {
       "type": "array",
       "items": [
        {
         "type": "object",
         "properties": {
          "indexFieldName": {
           "type": "string"
          },
          "indexFieldType": {
           "type": "string",
           "enum": [
            "STRING",
            "STRING_LIST",
            "DATE"
           ]
          },
          "dataSourceFieldName": {
           "type": "string"
          },
          "dateFieldFormat": {
           "type": "string",
           "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
          }
         },
         "required": [
          "indexFieldName",
          "indexFieldType",
          "dataSourceFieldName"
         ]
        }
       ]
      }
     },
     "required": [
      "fieldMappings"
     ]
    },
    "attachment": {
     "type": "object",
     "properties": {
      "fieldMappings": {
       "type": "array",
       "items": [
```

```
        {
         "type": "object",
         "properties": {
          "indexFieldName": {
           "type": "string"
          },
          "indexFieldType": {
           "type": "string",
           "enum": [
            "STRING",
            "STRING_LIST",
            "DATE"
           ]
          },
          "dataSourceFieldName": {
           "type": "string"
          },
          "dateFieldFormat": {
           "type": "string",
           "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
          }
         },
         "required": [
          "indexFieldName",
          "indexFieldType",
          "dataSourceFieldName"
         ]
        }
       ]
      }
     },
     "required": [
      "fieldMappings"
     ]
    },
    "comment": {
     "type": "object",
     "properties": {
      "fieldMappings": {
       "type": "array",
       "items": [
        {
         "type": "object",
         "properties": {
```

```
          "indexFieldName": {
           "type": "string"
          },
          "indexFieldType": {
           "type": "string",
           "enum": [
            "STRING",
            "STRING_LIST",
            "DATE"
           ]
          },
          "dataSourceFieldName": {
           "type": "string"
          },
          "dateFieldFormat": {
           "type": "string",
           "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
          }
         },
         "required": [
          "indexFieldName",
          "indexFieldType",
          "dataSourceFieldName"
         ]
        }
       ]
      }
     },
     "required": [
      "fieldMappings"
     ]
    }
   }
  },
  "additionalProperties": {
   "type": "object",
   "properties": {
    "eventTitleFilterRegEx": {
     "type": "array",
     "items": {
      "type": "string"
     }
    },
    "pageTitleFilterRegEx": {
```

```
      "type": "array",
      "items": {
       "type": "string"
      }
     },
     "linkTitleFilterRegEx": {
      "type": "array",
      "items": {
       "type": "string"
      }
     },
     "inclusionFilePath": {
      "type": "array",
      "items": {
       "type": "string"
      }
     },
     "exclusionFilePath": {
      "type": "array",
      "items": {
       "type": "string"
      }
     },
     "inclusionFileTypePatterns": {
      "type": "array",
      "items": {
       "type": "string"
      }
     },
     "exclusionFileTypePatterns": {
      "type": "array",
      "items": {
       "type": "string"
      }
     },
     "inclusionFileNamePatterns": {
      "type": "array",
      "items": {
       "type": "string"
      }
     },
     "exclusionFileNamePatterns": {
      "type": "array",
      "items": {
```

```
      "type": "string"
     }
    },
    "inclusionOneNoteSectionNamePatterns": {
     "type": "array",
     "items": {
      "type": "string"
     }
    },
    "exclusionOneNoteSectionNamePatterns": {
     "type": "array",
     "items": {
      "type": "string"
     }
    },
    "inclusionOneNotePageNamePatterns": {
     "type": "array",
     "items": {
      "type": "string"
     }
    },
    "exclusionOneNotePageNamePatterns": {
     "type": "array",
     "items": {
      "type": "string"
     }
    },
    "crawlFiles": {
     "type": "boolean"
    },
    "crawlPages": {
     "type": "boolean"
    },
    "crawlEvents": {
     "type": "boolean"
    },
    "crawlComments": {
     "type": "boolean"
    },
    "crawlLinks": {
     "type": "boolean"
    },
    "crawlAttachments": {
     "type": "boolean"
```

```
    },
    "crawlListData": {
     "type": "boolean"
    },
    "crawlAcl": {
     "type": "boolean"
    },
    "fieldForUserId": {
     "type": "string"
    },
    "aclConfiguration": {
     "type": "string",
     "enum": [
      "ACLWithLDAPEmailFmt",
      "ACLWithManualEmailFmt",
      "ACLWithUsernameFmt"
     ]
    },
    "emailDomain": {
     "type": "string"
    },
    "isCrawlLocalGroupMapping": {
     "type": "boolean"
    },
    "isCrawlAdGroupMapping": {
     "type": "boolean"
    },
    "proxyHost": {
     "type": "string"
    },
    "proxyPort": {
     "type": "string"
    }
   },
   "required": [
   ]
  },
  "type": {
   "type": "string",
   "pattern": "SHAREPOINTV2"
  },
  "enableIdentityCrawler": {
   "type": "boolean"
  },
```

```
  "syncMode": {
   "type": "string",
   "enum": [
    "FULL_CRAWL",
    "FORCED_FULL_CRAWL",
    "CHANGE_LOG"
   ]
  },
  "secretArn": {
   "type": "string",
   "minLength": 20,
   "maxLength": 2048
  }
 },
 "version": {
  "type": "string",
  "anyOf": [
   {
    "pattern": "1.0.0"
   }
  ]
 },
 "required": [
  "connectionConfiguration",
  "repositoryConfigurations",
  "enableIdentityCrawler",
  "syncMode",
  "additionalProperties",
  "secretArn",
  "type"
 ]
}
```

Schéma de modèle Microsoft SQL Server

Vous incluez un JSON qui contient le schéma de la source de données dans le cadre de l'TemplateConfigurationobjet. Spécifiez le type de source de donnéesJDBC, le type de base de donnéessqlserver, un secret pour vos informations d'authentification et les autres configurations nécessaires. Vous spécifiez ensuite TEMPLATE le Type moment de l'appel CreateDataSource.

Vous pouvez utiliser le modèle fourni dans ce guide du développeur. Consultez Schéma JSON de Microsoft SQL Server.

Le tableau suivant décrit les paramètres du schéma JSON de Microsoft SQL Server.

Configuration	Description
Configuration de connexion	Informations de configuration pour le point de terminaison de la source de données.
repositoryEndpointMetadata	Informations de configuration requises pour connecter votre source de données. • DBType : le type de base de données Java que vous utilisez, que ce soit, `mysql` `db2`, `postgresql` ou. `oracle` `sqlserver` • DBHost : nom d'hôte de la base de données. • DBPort : port de base de données. • DBInstance : instance de base de données.
Configurations du référentiel	Informations de configuration pour le contenu de la source de données. Par exemple, configurer des types spécifiques de contenu et des mappages de champs. Spécifiez le type de source de données et l'ARN secret.
document	Liste d'objets qui mappent les attributs ou les noms de champs du contenu de votre base de données aux noms de champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.
Propriétés supplémentaires	Options de configuration supplémentaires pour votre contenu dans votre source de données. À utiliser pour inclure ou exclure un contenu spécifique dans la source de données de votre base de données.

Configuration	Description
Clé primaire	Fournissez la clé primaire pour la table de base de données. Cela permet d'identifier une table au sein de votre base de données.
Colonne de titre	Indiquez le nom de la colonne du titre du document dans votre table de base de données.
Colonne Body	Indiquez le nom de la colonne du titre du document dans votre table de base de données.
sqlQuery	Entrez des instructions de requête SQL telles que les opérations SELECT et JOIN. La taille des requêtes SQL doit être inférieure à 32 Ko. Amazon Kendra analysera tout le contenu de la base de données correspondant à votre requête.
Colonne d'horodatage	Entrez le nom de la colonne contenant les horodatages. Amazon Kendra utilise les informations d'horodatage pour détecter les modifications apportées à votre contenu et synchroniser uniquement le contenu modifié.
Format d'horodatage	Entrez le nom de la colonne contenant les formats d'horodatage à utiliser pour détecter les modifications de contenu et resynchroniser votre contenu.
timezone	Entrez le nom de la colonne contenant les fuseaux horaires du contenu à analyser.

Configuration	Description
changeDetectingColumns	Entrez le nom des colonnes qui Amazon Kendra seront utilisées pour détecter les modifications de contenu. Amazon Kendra réindexera le contenu en cas de modification de l'une de ces colonnes
allowedUsersColumns	Entrez le nom de la colonne contenant l'utilisa teur devant IDs être autorisé à accéder au contenu.
allowedGroupsColumn	Entrez le nom de la colonne contenant l'utilisa teur devant IDs être autorisé à accéder au contenu.
source URIColumn	Entrez le nom de la colonne contenant la source URLs à indexer.
isSslEnabled	Entrez des instructions de requête SQL telles que les opérations SELECT et JOIN. La taille des requêtes SQL doit être inférieure à 32 Ko. Amazon Kendra analysera tout le contenu de la base de données correspondant à votre requête.
type	Type de source de données. Spécifiez `JDBC` comme type de source de données.

Configuration	Description
Mode de synchronisation	Spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Vous pouvez choisir entre : • `FORCED_FULL_CRAWL` pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index. • `FULL_CRAWL` pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modificat ions de contenu et indexer le contenu modifié depuis la dernière synchronisation. • `CHANGE_LOG` pour indexer uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchroni se avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

Configuration	Description
Secrétaire Arn	Le nom de ressource Amazon (ARN) d'un secret Secrets Manager contenant le nom d'utilisateur et le mot de passe requis pour se connecter à votre base de données. Le secret doit contenir une structure JSON avec les clés suivantes : `{ "user name": "database user name", "password": " password" }`
version	Version du modèle actuellement prise en charge.

Schéma JSON de Microsoft SQL Server

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "dbType": {
              "type": "string",
              "enum": [
                "mysql",
                "db2",
                "postgresql",
                "oracle",
                "sqlserver"
              ]
            },
            "dbHost": {
```

```
            "type": "string"
          },
          "dbPort": {
            "type": "string"
          },
          "dbInstance": {
            "type": "string"
          }
        },
        "required": [
          "dbType",
          "dbHost",
          "dbPort",
          "dbInstance"
        ]
      }
    },
    "required": [
      "repositoryEndpointMetadata"
    ]
  },
  "repositoryConfigurations": {
    "type": "object",
    "properties": {
      "document": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string"
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  }
                },
                "required": [
```

```
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      }
    },
    "required": [
    ]
  },
  "additionalProperties": {
    "type": "object",
    "properties": {
      "primaryKey": {
        "type": "string"
      },
      "titleColumn": {
        "type": "string"
      },
      "bodyColumn": {
        "type": "string"
      },
      "sqlQuery": {
        "type": "string",
        "not": {
          "pattern": ";+"
        }
      },
      "timestampColumn": {
        "type": "string"
      },
      "timestampFormat": {
        "type": "string"
      },
      "timezone": {
        "type": "string"
      },
      "changeDetectingColumns": {
```

```
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "allowedUsersColumn": {
          "type": "string"
        },
        "allowedGroupsColumn": {
          "type": "string"
        },
        "sourceURIColumn": {
          "type": "string"
        },
        "isSslEnabled": {
          "type": "boolean"
        }
      },
      "required": ["primaryKey", "titleColumn", "bodyColumn", "sqlQuery"]
    },
    "type" : {
      "type" : "string",
      "pattern": "JDBC"
    },
    "syncMode": {
      "type": "string",
      "enum": [
        "FORCED_FULL_CRAWL",
        "FULL_CRAWL",
        "CHANGE_LOG"
      ]
    },
    "secretArn": {
      "type": "string"
    }
  },
  "version": {
    "type": "string",
    "anyOf": [
      {
        "pattern": "1.0.0"
      }
    ]
  },
```

```
  "required": [
      "connectionConfiguration",
      "repositoryConfigurations",
      "syncMode",
      "additionalProperties",
      "secretArn",
      "type"
  ]
}
```

Schéma du modèle Microsoft Teams

Vous incluez un JSON qui contient le schéma de la source de données dans le cadre de l'TemplateConfigurationobjet. Vous fournissez l'ID du locataire dans le cadre de la configuration de la connexion ou des détails du point de terminaison du référentiel. Spécifiez également le type de source de donnéesMSTEAMS, un secret pour vos informations d'authentification et les autres configurations nécessaires. Vous spécifiez ensuite TEMPLATE le Type moment de l'appel CreateDataSource.

Vous pouvez utiliser le modèle fourni dans ce guide du développeur. Consultez Schéma Microsoft Teams JSON.

Le tableau suivant décrit les paramètres du schéma Microsoft Teams JSON.

Configuration	Description
Configuration de connexion	Informations de configuration du point de terminaison pour la source de données.
repositoryEndpointMetadata	Informations sur le point de terminaison de la source de données.
ID du locataire	L'identifiant du client Microsoft 365. Vous pouvez trouver votre ID de locataire dans les propriétés de votre portail Azure Active Directory ou dans votre OAuth application.
Configurations du référentiel	Informations de configuration pour le contenu de la source de données. Par exemple,

Configuration	Description
	configurer des types spécifiques de contenu et des mappages de champs.
• Message de chat • Pièce jointe au chat • ChannelPost • Wiki de la chaîne • Fixation du canal • Chat de réunion • Dossier de réunion • Note de réunion • Calendrier de la réunion	Liste d'objets qui mappent les attributs ou les noms de champs de votre contenu Microsoft Teams aux noms de champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.
Propriétés supplémentaires	Options de configuration supplémentaires pour votre contenu dans votre source de données.
Modèle de paiement	Spécifie le type de modèle de paiement à utiliser avec votre source de données Microsoft Teams. Les modèles de paiement du modèle A sont limités aux modèles de licence et de paiement qui nécessitent une conformité en matière de sécurité. Les modèles de paiement du modèle B conviennent aux modèles de licence et de paiement qui ne nécessitent pas de conformité en matière de sécurité.

Configuration	Description
• inclusionTeamNameFiltre • inclusionChannelNameFiltre • inclusionFileNameMotifs • inclusionFileTypeMotifs • inclusionUserEmailFiltre • inclusionOneNoteSectionNamePatterns • inclusionOneNotePageNamePatterns	Liste de modèles d'expressions régulières permettant d'inclure certains contenus dans votre source de données Microsoft Teams. Le contenu qui correspond aux modèles est inclus dans l'index. Le contenu qui ne correspond pas aux modèles est exclu de l'index. Si le contenu correspond à la fois à un modèle d'inclusion et à un modèle d'exclusion, le modèle d'exclusi on est prioritaire et le contenu n'est pas inclus dans l'index.
• exclusionTeamNameFiltre • exclusionChannelNameFiltre • exclusionFileNameMotifs • exclusionFileTypeMotifs • exclusionUserEmailFiltre • exclusionOneNoteSectionNamePatterns • exclusionOneNotePageNamePatterns	Liste de modèles d'expressions régulières permettant d'exclure certains contenus de votre source de données Microsoft Teams. Le contenu qui correspond aux modèles est exclu de l'index. Le contenu qui ne correspon d pas aux modèles est inclus dans l'index. Si le contenu correspond à la fois à un modèle d'inclusion et à un modèle d'exclusion, le modèle d'exclusion est prioritaire et le contenu n'est pas inclus dans l'index.
• isCrawlChatUn message • isCrawlChatPièce jointe • isCrawlChannelPublier • isCrawlChannelPièce jointe • isCrawlChannelWiki • isCrawlCalendarRéunion • isCrawlMeetingDiscuter • isCrawlMeetingDossier • isCrawlMeetingRemarque	`true`pour analyser ces types de contenu dans votre source de données Microsoft Teams.

Configuration	Description
startCalendarDateHeure	Vous pouvez configurer une date et une heure de début spécifiques pour le contenu de votre calendrier.
endCalendarDateHeure	Vous pouvez configurer une date et une heure de fin spécifiques pour le contenu du calendrier.
type	Type de source de données. Spécifiez `MSTEAMS` comme type de source de données.
enableIdentityCrawler	`true`utiliser Amazon Kendra le moteur de recherche d'identité pour synchroniser les identity/principal informations relatives aux utilisateurs et aux groupes ayant accès à certains documents. Si le robot d'identification est désactivé, tous les documents peuvent être consultés publiquement. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'identification est désactivé , vous pouvez également utiliser l'PutPrincipalMappingAPI pour télécharger les informations d'accès des utilisateurs et des groupes.

Configuration	Description
Mode de synchronisation	Spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Vous pouvez choisir entre : • `FORCED_FULL_CRAWL` pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index. • `FULL_CRAWL` pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modificat ions de contenu et indexer le contenu modifié depuis la dernière synchronisation. • `CHANGE_LOG` pour indexer uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchroni se avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
Secrétaire Arn	Le nom de ressource Amazon (ARN) d'un AWS Secrets Manager secret qui contient les paires clé-valeur requises pour se connecter à vos Microsoft Teams. Cela inclut votre ID client et le secret client générés lorsque vous créez une OAuth application sur le portail Azure.

Configuration	Description
version	Version de ce modèle actuellement prise en charge.

Schéma Microsoft Teams JSON

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "tenantId": {
              "type": "string",
              "pattern": "^[0-9a-f]{8}-[0-9a-f]{4}-[0-9a-f]{4}-[0-9a-f]{4}-[0-9a-f]
{12}$",
              "minLength": 36,
              "maxLength": 36
            }
          },
          "required": [
            "tenantId"
          ]
        }
      },
      "required": [
        "repositoryEndpointMetadata"
      ]
    },
    "repositoryConfigurations": {
      "type": "object",
      "properties": {
        "chatMessage": {
          "type": "object",
          "properties": {
            "fieldMappings": {
              "type": "array",
```

```
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string",
                    "enum": [
                      "STRING",
                      "STRING_LIST",
                      "DATE"
                    ]
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  },
                  "dateFieldFormat": {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required": [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      },
      "chatAttachment": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
```

```
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string",
                    "enum": [
                      "STRING",
                      "DATE",
                      "LONG"
                    ]
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  },
                  "dateFieldFormat": {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required": [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      },
      "channelPost": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
```

```
                },
                "indexFieldType": {
                  "type": "string",
                  "enum": [
                    "STRING",
                    "STRING_LIST",
                    "DATE"
                  ]
                },
                "dataSourceFieldName": {
                  "type": "string"
                },
                "dateFieldFormat": {
                  "type": "string",
                  "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                }
              },
              "required": [
                "indexFieldName",
                "indexFieldType",
                "dataSourceFieldName"
              ]
            }
          ]
        }
      },
      "required": [
        "fieldMappings"
      ]
    },
    "channelWiki": {
      "type": "object",
      "properties": {
        "fieldMappings": {
          "type": "array",
          "items": [
            {
              "type": "object",
              "properties": {
                "indexFieldName": {
                  "type": "string"
                },
                "indexFieldType": {
                  "type": "string",
```

```
                    "enum": [
                      "STRING",
                      "DATE",
                      "LONG"
                    ]
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  },
                  "dateFieldFormat": {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required": [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      },
      "channelAttachment": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string",
                    "enum": [
                      "STRING",
                      "DATE",
```

```
                    "LONG"
                  ]
                },
                "dataSourceFieldName": {
                  "type": "string"
                },
                "dateFieldFormat": {
                  "type": "string",
                  "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                }
              },
              "required": [
                "indexFieldName",
                "indexFieldType",
                "dataSourceFieldName"
              ]
            }
          ]
        }
      },
      "required": [
        "fieldMappings"
      ]
    },
    "meetingChat": {
      "type": "object",
      "properties": {
        "fieldMappings": {
          "type": "array",
          "items": [
            {
              "type": "object",
              "properties": {
                "indexFieldName": {
                  "type": "string"
                },
                "indexFieldType": {
                  "type": "string",
                  "enum": [
                    "STRING",
                    "STRING_LIST",
                    "DATE"
                  ]
                },
```

```
                "dataSourceFieldName": {
                  "type": "string"
                },
                "dateFieldFormat": {
                  "type": "string",
                  "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                }
              },
              "required": [
                "indexFieldName",
                "indexFieldType",
                "dataSourceFieldName"
              ]
            }
          ]
        }
      },
      "required": [
        "fieldMappings"
      ]
    },
    "meetingFile": {
      "type": "object",
      "properties": {
        "fieldMappings": {
          "type": "array",
          "items": [
            {
              "type": "object",
              "properties": {
                "indexFieldName": {
                  "type": "string"
                },
                "indexFieldType": {
                  "type": "string",
                  "enum": [
                    "STRING",
                    "DATE",
                    "LONG"
                  ]
                },
                "dataSourceFieldName": {
                  "type": "string"
                },
```

```
                "dateFieldFormat": {
                  "type": "string",
                  "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                }
              },
              "required": [
                "indexFieldName",
                "indexFieldType",
                "dataSourceFieldName"
              ]
            }
          ]
        }
      },
      "required": [
        "fieldMappings"
      ]
    },
    "meetingNote": {
      "type": "object",
      "properties": {
        "fieldMappings": {
          "type": "array",
          "items": [
            {
              "type": "object",
              "properties": {
                "indexFieldName": {
                  "type": "string"
                },
                "indexFieldType": {
                  "type": "string",
                  "enum": [
                    "STRING",
                    "DATE"
                  ]
                },
                "dataSourceFieldName": {
                  "type": "string"
                },
                "dateFieldFormat": {
                  "type": "string",
                  "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                }
```

```
                },
                "required": [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      },
      "calendarMeeting": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string",
                    "enum": [
                      "STRING",
                      "DATE"
                    ]
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  },
                  "dateFieldFormat": {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required": [
                  "indexFieldName",
                  "indexFieldType",
```

```
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      }
    }
  },
   "additionalProperties": {
    "type": "object",
    "properties": {
      "paymentModel": {
        "type": "string",
        "enum": [
          "A",
          "B",
          "Evaluation Mode"
        ]
      },
      "inclusionTeamNameFilter": {
        "type": "array",
        "items": {
          "type": "string"
        }
      },
      "exclusionTeamNameFilter": {
        "type": "array",
        "items": {
          "type": "string"
        }
      },
      "inclusionChannelNameFilter": {
        "type": "array",
        "items": {
          "type": "string"
        }
      },
      "exclusionChannelNameFilter": {
        "type": "array",
        "items": {
```

```
            "type": "string"
          }
        },
        "inclusionFileNamePatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "exclusionFileNamePatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "inclusionFileTypePatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "exclusionFileTypePatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "inclusionUserEmailFilter": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "inclusionOneNoteSectionNamePatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "exclusionOneNoteSectionNamePatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
```

```
        },
        "inclusionOneNotePageNamePatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "exclusionOneNotePageNamePatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "isCrawlChatMessage": {
          "type": "boolean"
        },
        "isCrawlChatAttachment": {
          "type": "boolean"
        },
        "isCrawlChannelPost": {
          "type": "boolean"
        },
        "isCrawlChannelAttachment": {
          "type": "boolean"
        },
        "isCrawlChannelWiki": {
          "type": "boolean"
        },
        "isCrawlCalendarMeeting": {
          "type": "boolean"
        },
        "isCrawlMeetingChat": {
          "type": "boolean"
        },
        "isCrawlMeetingFile": {
          "type": "boolean"
        },
        "isCrawlMeetingNote": {
          "type": "boolean"
        },
        "startCalendarDateTime": {
          "anyOf": [
            {
              "type": "string",
```

```
            "pattern": "^[0-9]{4}-[0-9]{2}-[0-9]{2}T[0-9]{2}:[0-9]{2}:[0-9]{2}Z$"
          },
          {
            "type": "string",
            "pattern": ""
          }
        ]
      },
      "endCalendarDateTime": {
        "anyOf": [
          {
            "type": "string",
            "pattern": "^[0-9]{4}-[0-9]{2}-[0-9]{2}T[0-9]{2}:[0-9]{2}:[0-9]{2}Z$"
          },
          {
            "type": "string",
            "pattern": ""
          }
        ]
      }
    },
    "required": []
  },
  "type": {
    "type": "string",
    "pattern": "MSTEAMS"
  },
  "enableIdentityCrawler": {
    "type": "boolean"
  },
  "syncMode": {
    "type": "string",
    "enum": [
      "FORCED_FULL_CRAWL",
      "FULL_CRAWL",
      "CHANGE_LOG"
    ]
  },
  "secretArn": {
    "type": "string",
    "minLength": 20,
    "maxLength": 2048
  }
},
```

```
  "version": {
    "type": "string",
    "anyOf": [
      {
        "pattern": "1.0.0"
      }
    ]
  },
  "required": [
    "connectionConfiguration",
    "repositoryConfigurations",
    "syncMode",
    "additionalProperties",
    "secretArn",
    "type"
  ]
}
```

Schéma du modèle Microsoft Yammer

Vous incluez un JSON qui contient le schéma de la source de données dans le cadre de TemplateConfigurationl'objet. Spécifiez le type de source de donnéesYAMMER, un secret pour vos informations d'authentification et les autres configurations nécessaires. Vous le spécifiez ensuite TEMPLATE comme type lorsque vous appelez CreateDataSource.

Vous pouvez utiliser le modèle fourni dans ce guide du développeur.

Le tableau suivant décrit les paramètres du schéma Microsoft Yammer JSON.

Configuration	Description
Configuration de connexion	Informations de configuration pour la source de données.
repositoryEndpointMetadata	Informations sur le point de terminaison de la source de données. Cette source de données ne spécifie pas de point de terminaison dansrepositoryEndpointMetadata . Les informations de connexion sont plutôt incluses dans un AWS Secrets Manager secret que vous fournissezsecretArn .

Configuration	Description
Configurations du référentiel	Informations de configuration pour le contenu de la source de données. Par exemple, configurer des types spécifiques de contenu et des mappages de champs.
• communauté • utilisateur • message • attachment	Liste d'objets qui associent les attributs ou les noms de champs du contenu Microsoft Yammer aux noms des champs d'index Amazon Kendra. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.
Propriétés supplémentaires	Options de configuration supplémentaires pour votre contenu dans votre source de données
Modèles d'inclusion	Liste de modèles d'expressions régulières permettant d'inclure certains fichiers dans votre source de données Microsoft Yammer. Les fichiers qui correspondent aux modèles sont inclus dans l'index. Les fichiers qui ne correspondent pas aux modèles sont exclus de l'index. Si un fichier correspond à la fois à un modèle d'inclusion et d'exclusion, le modèle d'exclusion a la priorité et le fichier n'est pas inclus dans l'index.

Configuration	Description
Schémas d'exclusion	Liste de modèles d'expressions régulières permettant d'exclure certains fichiers de votre source de données Microsoft Yammer. Les fichiers qui correspondent aux modèles sont exclus de l'index. Les fichiers qui ne correspon dent pas aux modèles sont inclus dans l'index. Si un fichier correspond à la fois à un modèle d'exclusion et à un modèle d'inclusion, le modèle d'exclusion est prioritaire et le fichier n'est pas inclus dans l'index.
Depuis Date	Vous pouvez choisir de configurer un `sinceDate` paramètre afin que le connecteu r Microsoft Yammer analyse le contenu en fonction d'un paramètre spécifique. `sinceDate`
communityNameFilter	Vous pouvez choisir d'indexer un contenu communautaire spécifique.
• isCrawlMessage • isCrawlAttachment • isCrawlPrivateUn message	`true`pour analyser les messages, les pièces jointes et les messages privés.
type	Spécifiez `YAMMER` comme type de source de données.
Secrétaire Arn	Le nom de ressource Amazon (ARN) d'un AWS Secrets Manager secret qui contient les paires clé-valeur requises pour se connecter à votre Microsoft Yammer. Cela inclut votre nom d'utilisateur et votre mot de passe Microsoft Yammer, ainsi que l'ID client et le secret client générés lorsque vous créez une OAuth application sur le portail Azure.

Configuration	Description
useChangeLog	`true`pour utiliser le journal des modifications de Microsoft Yammer afin de déterminer quels documents doivent être mis à jour dans l'index.
Mode de synchronisation	Spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Vous pouvez choisir entre : • `FORCED_FULL_CRAWL` pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index. • `FULL_CRAWL` pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modificat ions de contenu et indexer le contenu modifié depuis la dernière synchronisation. • `CHANGE_LOG` pour indexer uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchroni se avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

Configuration	Description
enableIdentityCrawler	`true`utiliser Amazon Kendra le moteur de recherche d'identité pour synchroniser les identity/principal informations relatives aux utilisateurs et aux groupes ayant accès à certains documents. Si le robot d'identification est désactivé, tous les documents peuvent être consultés publiquement. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'identification est désactivé , vous pouvez également utiliser l'PutPrincipalMappingAPI pour télécharger les informations d'accès des utilisateurs et des groupes.

Schéma Microsoft Yammer JSON

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
          }
        }
      },
      "required": [
        "repositoryEndpointMetadata"
      ]
    },
    "repositoryConfigurations": {
      "type": "object",
      "properties": {
        "community": {
          "type": "object",
          "properties": {
```

```
            "fieldMappings": {
              "type": "array",
              "items": {
                "anyOf": [
                  {
                    "type": "object",
                    "properties": {
                      "indexFieldName": {
                        "type": "string"
                      },
                      "indexFieldType": {
                        "type": "string",
                        "enum": [
                          "STRING",
                          "DATE"
                        ]
                      },
                      "dataSourceFieldName": {
                        "type": "string"
                      },
                      "dateFieldFormat": {
                        "type": "string",
                        "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                      }
                    },
                    "required": [
                      "indexFieldName",
                      "indexFieldType",
                      "dataSourceFieldName"
                    ]
                  }
                ]
              }
            }
          },
          "required": [
            "fieldMappings"
          ]
        },
        "user": {
          "type": "object",
          "properties": {
            "fieldMappings": {
              "type": "array",
```

```
            "items": {
              "anyOf": [
                {
                  "type": "object",
                  "properties": {
                    "indexFieldName": {
                      "type": "string"
                    },
                    "indexFieldType": {
                      "type": "string",
                      "enum": [
                        "STRING",
                        "DATE"
                      ]
                    },
                    "dataSourceFieldName": {
                      "type": "string"
                    },
                    "dateFieldFormat": {
                      "type": "string",
                      "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                    }
                  },
                  "required": [
                    "indexFieldName",
                    "indexFieldType",
                    "dataSourceFieldName"
                  ]
                }
              ]
            }
          }
        },
        "required": [
          "fieldMappings"
        ]
      },
      "message": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": {
              "anyOf": [
```

```
                {
                  "type": "object",
                  "properties": {
                    "indexFieldName": {
                      "type": "string"
                    },
                    "indexFieldType": {
                      "type": "string",
                      "enum": [
                        "STRING",
                        "DATE"
                      ]
                    },
                    "dataSourceFieldName": {
                      "type": "string"
                    },
                    "dateFieldFormat": {
                      "type": "string",
                      "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                    }
                  },
                  "required": [
                    "indexFieldName",
                    "indexFieldType",
                    "dataSourceFieldName"
                  ]
                }
              ]
            }
          }
        },
        "required": [
          "fieldMappings"
        ]
      },
      "attachment": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": {
              "anyOf": [
                {
                  "type": "object",
```

```
                  "properties": {
                    "indexFieldName": {
                      "type": "string"
                    },
                    "indexFieldType": {
                      "type": "string",
                      "enum": [
                        "STRING",
                        "DATE"
                      ]
                    },
                    "dataSourceFieldName": {
                      "type": "string"
                    },
                    "dateFieldFormat": {
                      "type": "string",
                      "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                    }
                  },
                  "required": [
                    "indexFieldName",
                    "indexFieldType",
                    "dataSourceFieldName"
                  ]
                }
              ]
            }
          }
        },
        "required": [
          "fieldMappings"
        ]
      }
    }
  },
  "additionalProperties": {
    "type": "object",
    "properties": {
      "inclusionPatterns": {
        "type": "array"
      },
      "exclusionPatterns": {
        "type": "array"
      },
```

```
        "sinceDate": {
          "type": "string",
          "pattern": "^(19|2[0-9])[0-9]{2}-(0[1-9]|1[012])-(0[1-9]|[12][0-9]|
3[01])T(0[0-9]|1[0-9]|2[0-3]):([0-5][0-9]):([0-5][0-9])((\\+|-)(0[0-9]|1[0-9]|2[0-3]):
([0-5][0-9]))?$"
        },
        "communityNameFilter": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "isCrawlMessage": {
          "type": "boolean"
        },
        "isCrawlAttachment": {
          "type": "boolean"
        },
        "isCrawlPrivateMessage": {
          "type": "boolean"
        }
      },
      "required": [
        "sinceDate"
      ]
    },
    "type": {
      "type": "string",
      "pattern": "YAMMER"
    },
    "secretArn": {
      "type": "string",
      "minLength": 20,
      "maxLength": 2048
    },
    "useChangeLog": {
      "type": "string",
      "enum": [
        "true",
        "false"
      ]
    },
    "syncMode": {
      "type": "string",
```

```
      "enum": [
        "FORCED_FULL_CRAWL",
        "FULL_CRAWL",
        "CHANGE_LOG"
      ]
    },
    "enableIdentityCrawler": {
      "type": "boolean"
    },
    "version": {
      "type": "string",
      "anyOf": [
        {
          "pattern": "1.0.0"
        }
      ]
    }
  },
  "required": [
    "connectionConfiguration",
    "repositoryConfigurations",
    "additionalProperties",
    "type",
    "secretArn",
    "syncMode"
  ]
}
```

Schéma du modèle MySQL

Vous incluez un JSON qui contient le schéma de la source de données dans le cadre de l'TemplateConfigurationobjet. Spécifiez le type de source de donnéesJDBC, le type de base de donnéesmysql, un secret pour vos informations d'authentification et les autres configurations nécessaires. Vous spécifiez ensuite TEMPLATE le Type moment de l'appel CreateDataSource.

Vous pouvez utiliser le modèle fourni dans ce guide du développeur. Consultez Schéma JSON MySQL.

Le tableau suivant décrit les paramètres du schéma JSON MySQL.

Configuration	Description
Configuration de connexion	Informations de configuration pour le point de terminaison de la source de données.
repositoryEndpointMetadata	Informations de configuration requises pour connecter votre source de données. • DBType : le type de base de données Java que vous utilisez, que ce soit, `mysql` `db2`, `postgresql` ou. `oracle` `sqlserver` • DBHost : nom d'hôte de la base de données. • DBPort : port de base de données. • DBInstance : instance de base de données.
Configurations du référentiel	Informations de configuration pour le contenu de la source de données. Par exemple, configurer des types spécifiques de contenu et des mappages de champs. Spécifiez le type de source de données et l'ARN secret.
document	Liste d'objets qui mappent les attributs ou les noms de champs du contenu de votre base de données aux noms de champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.
Propriétés supplémentaires	Options de configuration supplémentaires pour votre contenu dans votre source de données. À utiliser pour inclure ou exclure un contenu spécifique dans la source de données de votre base de données.
Clé primaire	Fournissez la clé primaire pour la table de base de données. Cela permet d'identifier une table au sein de votre base de données.

Configuration	Description
Colonne de titre	Indiquez le nom de la colonne du titre du document dans votre table de base de données.
Colonne Body	Indiquez le nom de la colonne du titre du document dans votre table de base de données.
sqlQuery	Entrez des instructions de requête SQL telles que les opérations SELECT et JOIN. La taille des requêtes SQL doit être inférieure à 32 Ko. Amazon Kendra analysera tout le contenu de la base de données correspondant à votre requête.
Colonne d'horodatage	Entrez le nom de la colonne contenant les horodatages. Amazon Kendra utilise les informations d'horodatage pour détecter les modifications apportées à votre contenu et synchroniser uniquement le contenu modifié.
Format d'horodatage	Entrez le nom de la colonne contenant les formats d'horodatage à utiliser pour détecter les modifications de contenu et resynchroniser votre contenu.
timezone	Entrez le nom de la colonne contenant les fuseaux horaires du contenu à analyser.
changeDetectingColumns	Entrez le nom des colonnes qui Amazon Kendra seront utilisées pour détecter les modifications de contenu. Amazon Kendra réindexera le contenu en cas de modification de l'une de ces colonnes

Configuration	Description
allowedUsersColumns	Entrez le nom de la colonne contenant l'utilisa teur devant IDs être autorisé à accéder au contenu.
allowedGroupsColumn	Entrez le nom de la colonne contenant l'utilisa teur devant IDs être autorisé à accéder au contenu.
source URIColumn	Entrez le nom de la colonne contenant la source URLs à indexer.
isSslEnabled	Entrez des instructions de requête SQL telles que les opérations SELECT et JOIN. La taille des requêtes SQL doit être inférieure à 32 Ko. Amazon Kendra analysera tout le contenu de la base de données correspondant à votre requête.
type	Type de source de données. Spécifiez `JDBC` comme type de source de données.

Configuration	Description
Mode de synchronisation	Spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Vous pouvez choisir entre : • `FORCED_FULL_CRAWL` pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index. • `FULL_CRAWL` pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modificat ions de contenu et indexer le contenu modifié depuis la dernière synchronisation. • `CHANGE_LOG` pour indexer uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchroni se avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

Configuration	Description
Secrétaire Arn	Le nom de ressource Amazon (ARN) d'un secret Secrets Manager contenant le nom d'utilisateur et le mot de passe requis pour se connecter à votre base de données. Le secret doit contenir une structure JSON avec les clés suivantes : `{ "user name": "database user name", "password": " password" }`
version	Version du modèle actuellement prise en charge.

Schéma JSON MySQL

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "dbType": {
              "type": "string",
              "enum": [
                "mysql",
                "db2",
                "postgresql",
                "oracle",
                "sqlserver"
              ]
            },
            "dbHost": {
```

```
            "type": "string"
          },
          "dbPort": {
            "type": "string"
          },
          "dbInstance": {
            "type": "string"
          }
        },
        "required": [
          "dbType",
          "dbHost",
          "dbPort",
          "dbInstance"
        ]
      }
    },
    "required": [
      "repositoryEndpointMetadata"
    ]
  },
  "repositoryConfigurations": {
    "type": "object",
    "properties": {
      "document": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string"
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  }
                },
                "required": [
```

```
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      }
    },
    "required": [
    ]
  },
  "additionalProperties": {
    "type": "object",
    "properties": {
      "primaryKey": {
        "type": "string"
      },
      "titleColumn": {
        "type": "string"
      },
      "bodyColumn": {
        "type": "string"
      },
      "sqlQuery": {
        "type": "string",
        "not": {
          "pattern": ";+"
        }
      },
      "timestampColumn": {
        "type": "string"
      },
      "timestampFormat": {
        "type": "string"
      },
      "timezone": {
        "type": "string"
      },
      "changeDetectingColumns": {
```

```
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "allowedUsersColumn": {
          "type": "string"
        },
        "allowedGroupsColumn": {
          "type": "string"
        },
        "sourceURIColumn": {
          "type": "string"
        },
        "isSslEnabled": {
          "type": "boolean"
        }
      },
      "required": ["primaryKey", "titleColumn", "bodyColumn", "sqlQuery"]
    },
    "type" : {
      "type" : "string",
      "pattern": "JDBC"
    },
    "syncMode": {
      "type": "string",
      "enum": [
        "FORCED_FULL_CRAWL",
        "FULL_CRAWL",
        "CHANGE_LOG"
      ]
    },
    "secretArn": {
      "type": "string"
    }
  },
  "version": {
    "type": "string",
    "anyOf": [
      {
        "pattern": "1.0.0"
      }
    ]
  },
```

```
  "required": [
      "connectionConfiguration",
      "repositoryConfigurations",
      "syncMode",
      "additionalProperties",
      "secretArn",
      "type"
  ]
}
```

Schéma du modèle de base de données Oracle

Vous incluez un JSON qui contient le schéma de la source de données dans le cadre de l'TemplateConfigurationobjet. Spécifiez le type de source de donnéesJDBC, le type de base de donnéesoracle, un secret pour vos informations d'authentification et les autres configurations nécessaires. Vous spécifiez ensuite TEMPLATE le Type moment de l'appel CreateDataSource.

Vous pouvez utiliser le modèle fourni dans ce guide du développeur. Consultez Schéma JSON de base de données Oracle.

Le tableau suivant décrit les paramètres du schéma JSON de la base de données Oracle.

Configuration	Description
Configuration de connexion	Informations de configuration pour le point de terminaison de la source de données.
repositoryEndpointMetadata	Informations de configuration requises pour connecter votre source de données. • DBType : le type de base de données Java que vous utilisez, que ce soit, mysqldb2, postgresql ou. oracle sqlserver • DBHost : nom d'hôte de la base de données. • DBPort : port de base de données. • DBInstance : instance de base de données.
Configurations du référentiel	Informations de configuration pour le contenu de la source de données. Par exemple,

Configuration	Description
	configurer des types spécifiques de contenu et des mappages de champs. Spécifiez le type de source de données et l'ARN secret.
document	Liste d'objets qui mappent les attributs ou les noms de champs du contenu de votre base de données aux noms de champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.
Propriétés supplémentaires	Options de configuration supplémentaires pour votre contenu dans votre source de données. À utiliser pour inclure ou exclure un contenu spécifique dans la source de données de votre base de données.
Clé primaire	Fournissez la clé primaire pour la table de base de données. Cela permet d'identifier une table au sein de votre base de données.
Colonne de titre	Indiquez le nom de la colonne du titre du document dans votre table de base de données.
Colonne Body	Indiquez le nom de la colonne du titre du document dans votre table de base de données.
sqlQuery	Entrez des instructions de requête SQL telles que les opérations SELECT et JOIN. La taille des requêtes SQL doit être inférieure à 32 Ko. Amazon Kendra analysera tout le contenu de la base de données correspondant à votre requête.

Configuration	Description
Colonne d'horodatage	Entrez le nom de la colonne contenant les horodatages. Amazon Kendra utilise les informations d'horodatage pour détecter les modifications apportées à votre contenu et synchroniser uniquement le contenu modifié.
Format d'horodatage	Entrez le nom de la colonne contenant les formats d'horodatage à utiliser pour détecter les modifications de contenu et resynchroniser votre contenu.
timezone	Entrez le nom de la colonne contenant les fuseaux horaires du contenu à analyser.
changeDetectingColumns	Entrez le nom des colonnes qui Amazon Kendra seront utilisées pour détecter les modifications de contenu. Amazon Kendra réindexera le contenu en cas de modification de l'une de ces colonnes
allowedUsersColumns	Entrez le nom de la colonne contenant l'utilisa teur devant IDs être autorisé à accéder au contenu.
allowedGroupsColumn	Entrez le nom de la colonne contenant l'utilisa teur devant IDs être autorisé à accéder au contenu.
source URIColumn	Entrez le nom de la colonne contenant la source URLs à indexer.

Configuration	Description
isSslEnabled	Entrez des instructions de requête SQL telles que les opérations SELECT et JOIN. La taille des requêtes SQL doit être inférieure à 32 Ko. Amazon Kendra analysera tout le contenu de la base de données correspondant à votre requête.
type	Type de source de données. Spécifiez `JDBC` comme type de source de données.

Configuration	Description
Mode de synchronisation	Spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Vous pouvez choisir entre : • `FORCED_FULL_CRAWL` pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index. • `FULL_CRAWL` pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation. • `CHANGE_LOG` pour indexer uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

Configuration	Description
Secrétaire Arn	Le nom de ressource Amazon (ARN) d'un secret Secrets Manager contenant le nom d'utilisateur et le mot de passe requis pour se connecter à votre base de données. Le secret doit contenir une structure JSON avec les clés suivantes : `{ "user name": "database user name", "password": " password" }`
version	Version du modèle actuellement prise en charge.

Schéma JSON de base de données Oracle

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "dbType": {
              "type": "string",
              "enum": [
                "mysql",
                "db2",
                "postgresql",
                "oracle",
                "sqlserver"
              ]
            },
            "dbHost": {
```

```
            "type": "string"
          },
          "dbPort": {
            "type": "string"
          },
          "dbInstance": {
            "type": "string"
          }
        },
        "required": [
          "dbType",
          "dbHost",
          "dbPort",
          "dbInstance"
        ]
      }
    },
    "required": [
      "repositoryEndpointMetadata"
    ]
  },
  "repositoryConfigurations": {
    "type": "object",
    "properties": {
      "document": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string"
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  }
                },
                "required": [
```

```
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      }
    },
    "required": [
    ]
  },
  "additionalProperties": {
    "type": "object",
    "properties": {
      "primaryKey": {
        "type": "string"
      },
      "titleColumn": {
        "type": "string"
      },
      "bodyColumn": {
        "type": "string"
      },
      "sqlQuery": {
        "type": "string",
        "not": {
          "pattern": ";+"
        }
      },
      "timestampColumn": {
        "type": "string"
      },
      "timestampFormat": {
        "type": "string"
      },
      "timezone": {
        "type": "string"
      },
      "changeDetectingColumns": {
```

```
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "allowedUsersColumn": {
          "type": "string"
        },
        "allowedGroupsColumn": {
          "type": "string"
        },
        "sourceURIColumn": {
          "type": "string"
        },
        "isSslEnabled": {
          "type": "boolean"
        }
      },
      "required": ["primaryKey", "titleColumn", "bodyColumn", "sqlQuery"]
    },
    "type" : {
      "type" : "string",
      "pattern": "JDBC"
    },
    "syncMode": {
      "type": "string",
      "enum": [
        "FORCED_FULL_CRAWL",
        "FULL_CRAWL",
        "CHANGE_LOG"
      ]
    },
    "secretArn": {
      "type": "string"
    }
  },
  "version": {
    "type": "string",
    "anyOf": [
      {
        "pattern": "1.0.0"
      }
    ]
  },
```

```
  "required": [
      "connectionConfiguration",
      "repositoryConfigurations",
      "syncMode",
      "additionalProperties",
      "secretArn",
      "type"
  ]
}
```

Schéma du modèle PostgreSQL

Vous incluez un JSON qui contient le schéma de la source de données dans le cadre de l'TemplateConfigurationobjet. Spécifiez le type de source de donnéesJDBC, le type de base de donnéespostgresql, un secret pour vos informations d'authentification et les autres configurations nécessaires. Vous spécifiez ensuite TEMPLATE le Type moment de l'appel CreateDataSource.

Vous pouvez utiliser le modèle fourni dans ce guide du développeur. Consultez Schéma JSON PostgreSQL.

Le tableau suivant décrit les paramètres du schéma JSON de PostgreSQL.

Configuration	Description
Configuration de connexion	Informations de configuration pour le point de terminaison de la source de données.
repositoryEndpointMetadata	Informations de configuration requises pour connecter votre source de données. • DBType : le type de base de données Java que vous utilisez, que ce soit, mysqldb2, postgresql ou. oracle sqlserver • DBHost : nom d'hôte de la base de données. • DBPort : port de base de données. • DBInstance : instance de base de données.
Configurations du référentiel	Informations de configuration pour le contenu de la source de données. Par exemple,

Configuration	Description
	configurer des types spécifiques de contenu et des mappages de champs. Spécifiez le type de source de données et l'ARN secret.
document	Liste d'objets qui mappent les attributs ou les noms de champs du contenu de votre base de données aux noms de champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.
Propriétés supplémentaires	Options de configuration supplémentaires pour votre contenu dans votre source de données. À utiliser pour inclure ou exclure un contenu spécifique dans la source de données de votre base de données.
Clé primaire	Fournissez la clé primaire pour la table de base de données. Cela permet d'identifier une table au sein de votre base de données.
Colonne de titre	Indiquez le nom de la colonne du titre du document dans votre table de base de données.
Colonne Body	Indiquez le nom de la colonne du titre du document dans votre table de base de données.
sqlQuery	Entrez des instructions de requête SQL telles que les opérations SELECT et JOIN. La taille des requêtes SQL doit être inférieure à 32 Ko. Amazon Kendra analysera tout le contenu de la base de données correspondant à votre requête.

Configuration	Description
Colonne d'horodatage	Entrez le nom de la colonne contenant les horodatages. Amazon Kendra utilise les informations d'horodatage pour détecter les modifications apportées à votre contenu et synchroniser uniquement le contenu modifié.
Format d'horodatage	Entrez le nom de la colonne contenant les formats d'horodatage à utiliser pour détecter les modifications de contenu et resynchroniser votre contenu.
timezone	Entrez le nom de la colonne contenant les fuseaux horaires du contenu à analyser.
changeDetectingColumns	Entrez le nom des colonnes qui Amazon Kendra seront utilisées pour détecter les modifications de contenu. Amazon Kendra réindexera le contenu en cas de modification de l'une de ces colonnes
allowedUsersColumns	Entrez le nom de la colonne contenant l'utilisa teur devant IDs être autorisé à accéder au contenu.
allowedGroupsColumn	Entrez le nom de la colonne contenant l'utilisa teur devant IDs être autorisé à accéder au contenu.
source URIColumn	Entrez le nom de la colonne contenant la source URLs à indexer.

Configuration	Description
isSslEnabled	Entrez des instructions de requête SQL telles que les opérations SELECT et JOIN. La taille des requêtes SQL doit être inférieure à 32 Ko. Amazon Kendra analysera tout le contenu de la base de données correspondant à votre requête.
type	Type de source de données. Spécifiez `JDBC` comme type de source de données.

Configuration	Description
Mode de synchronisation	Spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Vous pouvez choisir entre : • `FORCED_FULL_CRAWL` pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index. • `FULL_CRAWL` pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modificat ions de contenu et indexer le contenu modifié depuis la dernière synchronisation. • `CHANGE_LOG` pour indexer uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchroni se avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

Configuration	Description
Secrétaire Arn	Le nom de ressource Amazon (ARN) d'un secret Secrets Manager contenant le nom d'utilisateur et le mot de passe requis pour se connecter à votre base de données. Le secret doit contenir une structure JSON avec les clés suivantes : `{ "user name": "database user name", "password": " password" }`
version	Version du modèle actuellement prise en charge.

Schéma JSON PostgreSQL

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "dbType": {
              "type": "string",
              "enum": [
                "mysql",
                "db2",
                "postgresql",
                "oracle",
                "sqlserver"
              ]
            },
            "dbHost": {
```

```
            "type": "string"
          },
          "dbPort": {
            "type": "string"
          },
          "dbInstance": {
            "type": "string"
          }
        },
        "required": [
          "dbType",
          "dbHost",
          "dbPort",
          "dbInstance"
        ]
      }
    },
    "required": [
      "repositoryEndpointMetadata"
    ]
  },
  "repositoryConfigurations": {
    "type": "object",
    "properties": {
      "document": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string"
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  }
                },
                "required": [
```

```
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      }
    },
    "required": [
    ]
  },
  "additionalProperties": {
    "type": "object",
    "properties": {
      "primaryKey": {
        "type": "string"
      },
      "titleColumn": {
        "type": "string"
      },
      "bodyColumn": {
        "type": "string"
      },
      "sqlQuery": {
        "type": "string",
        "not": {
          "pattern": ";+"
        }
      },
      "timestampColumn": {
        "type": "string"
      },
      "timestampFormat": {
        "type": "string"
      },
      "timezone": {
        "type": "string"
      },
      "changeDetectingColumns": {
```

```
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "allowedUsersColumn": {
          "type": "string"
        },
        "allowedGroupsColumn": {
          "type": "string"
        },
        "sourceURIColumn": {
          "type": "string"
        },
        "isSslEnabled": {
          "type": "boolean"
        }
      },
      "required": ["primaryKey", "titleColumn", "bodyColumn", "sqlQuery"]
    },
    "type" : {
      "type" : "string",
      "pattern": "JDBC"
    },
    "syncMode": {
      "type": "string",
      "enum": [
        "FORCED_FULL_CRAWL",
        "FULL_CRAWL",
        "CHANGE_LOG"
      ]
    },
    "secretArn": {
      "type": "string"
    }
  },
  "version": {
    "type": "string",
    "anyOf": [
      {
        "pattern": "1.0.0"
      }
    ]
  },
```

```
  "required": [
      "connectionConfiguration",
      "repositoryConfigurations",
      "syncMode",
      "additionalProperties",
      "secretArn",
      "type"
  ]
}
```

Schéma du modèle Salesforce

Vous incluez un JSON qui contient le schéma de la source de données dans le cadre de l'TemplateConfigurationobjet. Vous fournissez l'URL de l'hôte Salesforce dans le cadre de la configuration de la connexion ou des détails du point de terminaison du référentiel. Spécifiez également le type de source de donnéesSALESFORCEV2, un secret pour vos informations d'authentification et les autres configurations nécessaires. Vous spécifiez ensuite TEMPLATE le Type moment de l'appel CreateDataSource.

Vous pouvez utiliser le modèle fourni dans ce guide du développeur. Consultez Schéma JSON Salesforce.

Le tableau suivant décrit les paramètres du schéma JSON Salesforce.

Configuration	Description
Configuration de connexion	Informations de configuration pour le point de terminaison de la source de données.
repositoryEndpointMetadata	Informations sur le point de terminaison de la source de données.
URL de l'hôte	URL de l'instance Salesforce à indexer.
Configurations du référentiel	Informations de configuration pour le contenu de la source de données. Par exemple, configurer des types spécifiques de contenu et des mappages de champs.
• compte	Liste d'objets qui mappent les attributs ou les noms de champs de vos entités Salesforce

Configuration	Description
• contact • campaign • cas • produit • lead • contrat • partenaire • profile • idée • livre de prix • tâche • solution • attachment • utilisateur • document • Articles de connaissances • groupe • opportunité • bavarder • Entité personnalisée	aux noms de champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.

Configuration	Description
Secrétaire Arn	Le nom de ressource Amazon (ARN) d'un AWS Secrets Manager secret qui contient les paires clé-valeur requises pour se connecter à votre Salesforce. Le secret doit contenir une structure JSON avec les clés suivantes : `{ "authenticationUrl": " OAUTH endpoint that Amazon Kendra connects to get an OAUTH token", "consumerKey": " Application public key generated when you created your Salesforce application ", "consumerSecret": " Application private key generated when you created your Salesforce application ", "password": " Password associated with the user logging in to the Salesforce instance ", "securityToken": " Token associated with the user account logging in to the Salesforce instance ", "username": " User name of the user logging in to the Salesforce instance" }`
Propriétés supplémentaires	Options de configuration supplémentaires pour votre contenu dans votre source de données

Configuration	Description
• Filtre de compte • Filtre de contact • Filtre Case • Filtre de campagne • Filtre de contrat • Filtre de groupe • Filtre au plomb • Filtre de produit • Filtre d'opportunité • Filtre pour les partenaires • Filtre PriceBook • Filtre Idea • Filtre de profil • Filtre de tâches • Filtre de solution • Filtre utilisateur • Filtre Chatter • Filtre de documents • knowledgeArticleFilter • Entités personnalisées	Collection de chaînes qui indique les entités à filtrer.

Configuration	Description
Modèles d'inclusion • inclusionDocumentFileTypePatterns • inclusionDocumentFileNamePatterns • inclusionAccountFileTypePatterns • inclusionCampaignFileTypePatterns • inclusionDocumentFileNamePatterns • inclusionCampaignFileNamePatterns • inclusionCaseFileTypePatterns • inclusionCaseFileNamePatterns • inclusionContactFileTypePatterns • inclusionContractFileNamePatterns • inclusionLeadFileTypePatterns • inclusionLeadFileNamePatterns • inclusionOpportunityFileTypePatterns • inclusionOpportunityFileNamePatterns • inclusionSolutionFileTypePatterns • inclusionSolutionFileNamePatterns • inclusionTaskFileTypePatterns • inclusionTaskFileNamePatterns • inclusionGroupFileTypePatterns • inclusionGroupFileNamePatterns • inclusionChatterFileTypePatterns • inclusionChatterFileNamePatterns • inclusionCustomEntityFileTypePatterns • inclusionCustomEntityFileNamePatterns	Liste de modèles d'expressions régulières permettant d'inclure certains fichiers dans votre source de données Salesforce. Les fichiers qui correspondent aux modèles sont inclus dans l'index. Les fichiers qui ne correspondent pas aux modèles sont exclus de l'index. Si un fichier correspond à la fois à un modèle d'inclusion et d'exclusion, le modèle d'exclusion a la priorité et le fichier n'est pas inclus dans l'index.

Configuration	Description
Schémas d'exclusion • exclusionDocumentFileTypePatterns • exclusionDocumentFileNamePatterns • exclusionAccountFileTypePatterns • exclusionCampaignFileTypePatterns • exclusionCampaignFileNamePatterns • exclusionCaseFileTypePatterns • exclusionCaseFileNamePatterns • exclusionContactFileTypePatterns • exclusionContractFileNamePatterns • exclusionLeadFileTypePatterns • exclusionLeadFileNamePatterns • exclusionOpportunityFileTypePatterns • exclusionOpportunityFileNamePatterns • exclusionSolutionFileTypePatterns • exclusionSolutionFileNamePatterns • exclusionTaskFileTypePatterns • exclusionTaskFileNamePatterns • exclusionGroupFileTypePatterns • exclusionGroupFileNamePatterns • exclusionChatterFileTypePatterns • exclusionChatterFileNamePatterns • exclusionCustomEntityFileTypePatterns • exclusionCustomEntityFileNamePatterns	Liste de modèles d'expressions régulières permettant d'exclure certains fichiers de votre source de données Salesforce. Les fichiers qui correspondent aux modèles sont exclus de l'index. Les fichiers qui ne correspondent pas aux modèles sont inclus dans l'index. Si un fichier correspond à la fois à un modèle d'exclusion et à un modèle d'inclusion, le modèle d'exclusion est prioritaire et le fichier n'est pas inclus dans l'index.

Configuration	Description
• isCrawlAccount • isCrawlContact • isCrawlCase • isCrawlCampaign • isCrawlProduct • isCrawlLead • isCrawlContract • isCrawlPartner • isCrawlProfile • isCrawlIdea • isCrawlPricebook • isCrawlDocument • crawlSharedDocument • isCrawlGroup • isCrawlOpportunity • isCrawlChatter • isCrawlUser • isCrawlSolution • isCrawlTask • isCrawlAccountPièces jointes • isCrawlContactPièces jointes • isCrawlCasePièces jointes • isCrawlCampaignPièces jointes • isCrawlLeadPièces jointes • isCrawlContractPièces jointes • isCrawlGroupPièces jointes • isCrawlOpportunityPièces jointes • isCrawlChatterPièces jointes • isCrawlSolutionPièces jointes	`true`pour explorer ces types de fichiers dans votre compte Salesforce.

Configuration	Description

• isCrawlTaskPièces jointes
• isCrawlCustomEntityAttachments
• isCrawlKnowledgeDes articles
  • isCrawlDraft
  • isCrawlPublish
  • isCrawlArchived	

type	Type de source de données. Spécifiez `SALESFORCEV2` comme type de source de données.
enableIdentityCrawler	`true`utiliser Amazon Kendra le moteur de recherche d'identité pour synchroniser les identity/principal informations relatives aux utilisateurs et aux groupes ayant accès à certains documents. Si le robot d'identification est désactivé, tous les documents peuvent être consultés publiquement. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'identification est désactivé , vous pouvez également utiliser l'PutPrincipalMappingAPI pour télécharger les informations d'accès des utilisateurs et des groupes.

Configuration	Description
Mode de synchronisation	Spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Vous pouvez choisir entre : • `FORCED_FULL_CRAWL` pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index. • `FULL_CRAWL` pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modificat ions de contenu et indexer le contenu modifié depuis la dernière synchronisation. • `CHANGE_LOG` pour indexer uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchroni se avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
version	Version de ce modèle actuellement prise en charge.

Schéma JSON Salesforce

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
```

```
  "properties":
  {
    "connectionConfiguration": {
      "type": "object",
      "properties":
      {
        "repositoryEndpointMetadata":
        {
          "type": "object",
          "properties":
          {
            "hostUrl":
            {
              "type": "string",
              "pattern": "https:.*"
            }
          },
          "required":
          [
            "hostUrl"
          ]
        }
      },
      "required":
      [
        "repositoryEndpointMetadata"
      ]
    },
    "repositoryConfigurations": {
      "type": "object",
      "properties":
      {
        "account":
        {
          "type": "object",
          "properties":
          {
            "fieldMappings":
            {
              "type": "array",
              "items":
              [
                {
                  "type": "object",
```

```
                "properties":
                {
                  "indexFieldName":
                  {
                    "type": "string"
                  },
                  "indexFieldType":
                  {
                    "type": "string",
                    "enum":
                    [
                      "STRING",
                      "STRING_LIST",
                      "DATE",
                      "LONG"
                    ]
                  },
                  "dataSourceFieldName":
                  {
                    "type": "string"
                  },
                  "dateFieldFormat":
                  {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required":
                [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required":
        [
          "fieldMappings"
        ]
      },
      "contact":
      {
```

```
          "type": "object",
          "properties":
          {
            "fieldMappings":
            {
              "type": "array",
              "items":
              [
                {
                  "type": "object",
                  "properties":
                  {
                    "indexFieldName":
                    {
                      "type": "string"
                    },
                    "indexFieldType":
                    {
                      "type": "string",
                      "enum":
                      [
                        "STRING",
                        "STRING_LIST",
                        "DATE"
                      ]
                    },
                    "dataSourceFieldName":
                    {
                      "type": "string"
                    },
                    "dateFieldFormat":
                    {
                      "type": "string",
                      "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                    }
                  },
                  "required":
                  [
                    "indexFieldName",
                    "indexFieldType",
                    "dataSourceFieldName"
                  ]
                }
              ]
```

```
            }
          },
          "required":
          [
            "fieldMappings"
          ]
        },
        "campaign":
        {
          "type": "object",
          "properties":
          {
            "fieldMappings":
            {
              "type": "array",
              "items":
              [
                {
                  "type": "object",
                  "properties":
                  {
                    "indexFieldName":
                    {
                      "type": "string"
                    },
                    "indexFieldType":
                    {
                      "type": "string",
                      "enum":
                      [
                        "STRING",
                        "STRING_LIST",
                        "DATE",
                        "LONG"
                      ]
                    },
                    "dataSourceFieldName":
                    {
                      "type": "string"
                    },
                    "dateFieldFormat":
                    {
                      "type": "string",
                      "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
```

```
                    }
                  },
                  "required":
                  [
                    "indexFieldName",
                    "indexFieldType",
                    "dataSourceFieldName"
                  ]
                }
              ]
            }
          },
          "required":
          [
            "fieldMappings"
          ]
        },
        "case":
        {
          "type": "object",
          "properties":
          {
            "fieldMappings":
            {
              "type": "array",
              "items":
              [
                {
                  "type": "object",
                  "properties":
                  {
                    "indexFieldName":
                    {
                      "type": "string"
                    },
                    "indexFieldType":
                    {
                      "type": "string",
                      "enum":
                      [
                        "STRING",
                        "STRING_LIST",
                        "DATE"
                      ]
```

```
                },
                "dataSourceFieldName":
                {
                  "type": "string"
                },
                "dateFieldFormat":
                {
                  "type": "string",
                  "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                }
              },
              "required":
              [
                "indexFieldName",
                "indexFieldType",
                "dataSourceFieldName"
              ]
            }
          ]
        }
      },
      "required":
      [
        "fieldMappings"
      ]
    },
    "product":
    {
      "type": "object",
      "properties":
      {
        "fieldMappings":
        {
          "type": "array",
          "items":
          [
            {
              "type": "object",
              "properties":
              {
                "indexFieldName":
                {
                  "type": "string"
                },
```

```
                  "indexFieldType":
                  {
                    "type": "string",
                    "enum":
                    [
                      "STRING",
                      "STRING_LIST",
                      "DATE"
                    ]
                  },
                  "dataSourceFieldName":
                  {
                    "type": "string"
                  },
                  "dateFieldFormat":
                  {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required":
                [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required":
        [
          "fieldMappings"
        ]
      },
      "lead":
      {
        "type": "object",
        "properties":
        {
          "fieldMappings":
          {
            "type": "array",
            "items":
```

```
            [
              {
                "type": "object",
                "properties":
                {
                  "indexFieldName":
                  {
                    "type": "string"
                  },
                  "indexFieldType":
                  {
                    "type": "string",
                    "enum":
                    [
                      "STRING",
                      "STRING_LIST",
                      "DATE",
                      "LONG"
                    ]
                  },
                  "dataSourceFieldName":
                  {
                    "type": "string"
                  },
                  "dateFieldFormat":
                  {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required":
                [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required":
        [
          "fieldMappings"
        ]
```

```
        },
        "contract":
        {
          "type": "object",
          "properties":
          {
            "fieldMappings":
            {
              "type": "array",
              "items":
              [
                {
                  "type": "object",
                  "properties":
                  {
                    "indexFieldName":
                    {
                      "type": "string"
                    },
                    "indexFieldType":
                    {
                      "type": "string",
                      "enum":
                      [
                        "STRING",
                        "STRING_LIST",
                        "DATE"
                      ]
                    },
                    "dataSourceFieldName":
                    {
                      "type": "string"
                    },
                    "dateFieldFormat":
                    {
                      "type": "string",
                      "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                    }
                  },
                  "required":
                  [
                    "indexFieldName",
                    "indexFieldType",
                    "dataSourceFieldName"
```

```
                ]
              }
            ]
          }
        },
        "required":
        [
          "fieldMappings"
        ]
      },
      "partner":
      {
        "type": "object",
        "properties":
        {
          "fieldMappings":
          {
            "type": "array",
            "items":
            [
              {
                "type": "object",
                "properties":
                {
                  "indexFieldName":
                  {
                    "type": "string"
                  },
                  "indexFieldType":
                  {
                    "type": "string",
                    "enum":
                    [
                      "STRING",
                      "STRING_LIST",
                      "DATE"
                    ]
                  },
                  "dataSourceFieldName":
                  {
                    "type": "string"
                  },
                  "dateFieldFormat":
                  {
```

```
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required":
                [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required":
        [
          "fieldMappings"
        ]
      },
      "profile":
      {
        "type": "object",
        "properties":
        {
          "fieldMappings":
          {
            "type": "array",
            "items":
            [
              {
                "type": "object",
                "properties":
                {
                  "indexFieldName":
                  {
                    "type": "string"
                  },
                  "indexFieldType":
                  {
                    "type": "string",
                    "enum":
                    [
                      "STRING",
                      "STRING_LIST",
```

```
                    "DATE"
                  ]
                },
                "dataSourceFieldName":
                {
                  "type": "string"
                },
                "dateFieldFormat":
                {
                  "type": "string",
                  "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                }
              },
              "required":
              [
                "indexFieldName",
                "indexFieldType",
                "dataSourceFieldName"
              ]
            }
          ]
        }
      },
      "required":
      [
        "fieldMappings"
      ]
    },
    "idea":
    {
      "type": "object",
      "properties":
      {
        "fieldMappings":
        {
          "type": "array",
          "items":
          [
            {
              "type": "object",
              "properties":
              {
                "indexFieldName":
                {
```

```
                    "type": "string"
                  },
                  "indexFieldType":
                  {
                    "type": "string",
                    "enum":
                    [
                      "STRING",
                      "STRING_LIST",
                      "DATE",
                      "LONG"
                    ]
                  },
                  "dataSourceFieldName":
                  {
                    "type": "string"
                  },
                  "dateFieldFormat":
                  {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required":
                [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required":
        [
          "fieldMappings"
        ]
      },
      "pricebook":
      {
        "type": "object",
        "properties":
        {
          "fieldMappings":
```

```
            {
              "type": "array",
              "items":
              [
                {
                  "type": "object",
                  "properties":
                  {
                    "indexFieldName":
                    {
                      "type": "string"
                    },
                    "indexFieldType":
                    {
                      "type": "string",
                      "enum":
                      [
                        "STRING",
                        "STRING_LIST",
                        "DATE"
                      ]
                    },
                    "dataSourceFieldName":
                    {
                      "type": "string"
                    },
                    "dateFieldFormat":
                    {
                      "type": "string",
                      "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                    }
                  },
                  "required":
                  [
                    "indexFieldName",
                    "indexFieldType",
                    "dataSourceFieldName"
                  ]
                }
              ]
            }
          },
          "required":
          [
```

```
            "fieldMappings"
          ]
        },
        "task":
        {
          "type": "object",
          "properties":
          {
            "fieldMappings":
            {
              "type": "array",
              "items":
              [
                {
                  "type": "object",
                  "properties":
                  {
                    "indexFieldName":
                    {
                      "type": "string"
                    },
                    "indexFieldType":
                    {
                      "type": "string",
                      "enum":
                      [
                        "STRING",
                        "STRING_LIST",
                        "DATE"
                      ]
                    },
                    "dataSourceFieldName":
                    {
                      "type": "string"
                    },
                    "dateFieldFormat":
                    {
                      "type": "string",
                      "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                    }
                  },
                  "required":
                  [
                    "indexFieldName",
```

```
                "indexFieldType",
                "dataSourceFieldName"
              ]
            }
          ]
        }
      },
      "required":
      [
        "fieldMappings"
      ]
    },
    "solution":
    {
      "type": "object",
      "properties":
      {
        "fieldMappings":
        {
          "type": "array",
          "items":
          [
            {
              "type": "object",
              "properties":
              {
                "indexFieldName":
                {
                  "type": "string"
                },
                "indexFieldType":
                {
                  "type": "string",
                  "enum":
                  [
                    "STRING",
                    "STRING_LIST",
                    "DATE"
                  ]
                },
                "dataSourceFieldName":
                {
                  "type": "string"
                },
```

```
                "dateFieldFormat":
                {
                  "type": "string",
                  "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                }
              },
              "required":
              [
                "indexFieldName",
                "indexFieldType",
                "dataSourceFieldName"
              ]
            }
          ]
        }
      },
      "required":
      [
        "fieldMappings"
      ]
    },
    "attachment":
    {
      "type": "object",
      "properties":
      {
        "fieldMappings":
        {
          "type": "array",
          "items":
          [
            {
              "type": "object",
              "properties":
              {
                "indexFieldName":
                {
                  "type": "string"
                },
                "indexFieldType":
                {
                  "type": "string",
                  "enum":
                  [
```

```
                    "STRING",
                    "STRING_LIST",
                    "DATE",
                    "LONG"
                  ]
                },
                "dataSourceFieldName":
                {
                  "type": "string"
                },
                "dateFieldFormat":
                {
                  "type": "string",
                  "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                }
              },
              "required":
              [
                "indexFieldName",
                "indexFieldType",
                "dataSourceFieldName"
              ]
            }
          ]
        }
      },
      "required":
      [
        "fieldMappings"
      ]
    },
    "user":
    {
      "type": "object",
      "properties":
      {
        "fieldMappings":
        {
          "type": "array",
          "items":
          [
            {
              "type": "object",
              "properties":
```

```
                {
                  "indexFieldName":
                  {
                    "type": "string"
                  },
                  "indexFieldType":
                  {
                    "type": "string",
                    "enum":
                    [
                      "STRING",
                      "STRING_LIST",
                      "DATE"
                    ]
                  },
                  "dataSourceFieldName":
                  {
                    "type": "string"
                  },
                  "dateFieldFormat":
                  {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required":
                [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required":
        [
          "fieldMappings"
        ]
      },
      "document":
      {
        "type": "object",
        "properties":
```

```
          {
            "fieldMappings":
            {
              "type": "array",
              "items":
              [
                {
                  "type": "object",
                  "properties":
                  {
                    "indexFieldName":
                    {
                      "type": "string"
                    },
                    "indexFieldType":
                    {
                      "type": "string",
                      "enum":
                      [
                        "STRING",
                        "STRING_LIST",
                        "DATE",
                        "LONG"
                      ]
                    },
                    "dataSourceFieldName":
                    {
                      "type": "string"
                    },
                    "dateFieldFormat":
                    {
                      "type": "string",
                      "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                    }
                  },
                  "required":
                  [
                    "indexFieldName",
                    "indexFieldType",
                    "dataSourceFieldName"
                  ]
                }
              ]
            }
```

```
      },
      "required":
      [
        "fieldMappings"
      ]
    },
    "knowledgeArticles":
    {
      "type": "object",
      "properties":
      {
        "fieldMappings":
        {
          "type": "array",
          "items":
          [
            {
              "type": "object",
              "properties":
              {
                "indexFieldName":
                {
                  "type": "string"
                },
                "indexFieldType":
                {
                  "type": "string",
                  "enum":
                  [
                    "STRING",
                    "STRING_LIST",
                    "DATE"
                  ]
                },
                "dataSourceFieldName":
                {
                  "type": "string"
                },
                "dateFieldFormat":
                {
                  "type": "string",
                  "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                }
              },
```

```
                "required":
                [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required":
        [
          "fieldMappings"
        ]
      },
      "group":
      {
        "type": "object",
        "properties":
        {
          "fieldMappings":
          {
            "type": "array",
            "items":
            [
              {
                "type": "object",
                "properties":
                {
                  "indexFieldName":
                  {
                    "type": "string"
                  },
                  "indexFieldType":
                  {
                    "type": "string",
                    "enum":
                    [
                      "STRING",
                      "STRING_LIST",
                      "DATE"
                    ]
                  },
                  "dataSourceFieldName":
```

```
                {
                  "type": "string"
                },
                "dateFieldFormat":
                {
                  "type": "string",
                  "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                }
              },
              "required":
              [
                "indexFieldName",
                "indexFieldType",
                "dataSourceFieldName"
              ]
            }
          ]
        }
      },
      "required":
      [
        "fieldMappings"
      ]
    },
    "opportunity":
    {
      "type": "object",
      "properties":
      {
        "fieldMappings":
        {
          "type": "array",
          "items":
          [
            {
              "type": "object",
              "properties":
              {
                "indexFieldName":
                {
                  "type": "string"
                },
                "indexFieldType":
                {
```

```
                    "type": "string",
                    "enum":
                    [
                      "STRING",
                      "STRING_LIST",
                      "DATE",
                      "LONG"
                    ]
                  },
                  "dataSourceFieldName":
                  {
                    "type": "string"
                  },
                  "dateFieldFormat":
                  {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required":
                [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required":
        [
          "fieldMappings"
        ]
      },
      "chatter":
      {
        "type": "object",
        "properties":
        {
          "fieldMappings":
          {
            "type": "array",
            "items":
            [
```

```
                {
                  "type": "object",
                  "properties":
                  {
                    "indexFieldName":
                    {
                      "type": "string"
                    },
                    "indexFieldType":
                    {
                      "type": "string",
                      "enum":
                      [
                        "STRING",
                        "STRING_LIST",
                        "DATE"
                      ]
                    },
                    "dataSourceFieldName":
                    {
                      "type": "string"
                    },
                    "dateFieldFormat":
                    {
                      "type": "string",
                      "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                    }
                  },
                  "required":
                  [
                    "indexFieldName",
                    "indexFieldType",
                    "dataSourceFieldName"
                  ]
                }
              ]
            }
          },
          "required":
          [
            "fieldMappings"
          ]
        },
        "customEntity":
```

```
        {
          "type": "object",
          "properties":
          {
            "fieldMappings":
            {
              "type": "array",
              "items":
              [
                {
                  "type": "object",
                  "properties":
                  {
                    "indexFieldName":
                    {
                      "type": "string"
                    },
                    "indexFieldType":
                    {
                      "type": "string",
                      "enum":
                      [
                        "STRING",
                        "STRING_LIST",
                        "DATE"
                      ]
                    },
                    "dataSourceFieldName":
                    {
                      "type": "string"
                    },
                    "dateFieldFormat":
                    {
                      "type": "string",
                      "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                    }
                  },
                  "required":
                  [
                    "indexFieldName",
                    "indexFieldType",
                    "dataSourceFieldName"
                  ]
                }
```

```
            ]
          }
        },
        "required":
        [
          "fieldMappings"
        ]
      }
    }
  },
  "additionalProperties": {
    "type": "object",
    "properties":
    {
      "accountFilter":{
        "type": "array",
        "items":
        {
          "type": "string"
        }
      },
      "contactFilter":{
        "type": "array",
        "items":
        {
          "type": "string"
        }
      },
      "caseFilter":{
        "type": "array",
        "items":
        {
          "type": "string"
        }
      },
      "campaignFilter":{
        "type": "array",
        "items":
        {
          "type": "string"
        }
      },
      "contractFilter":{
        "type": "array",
```

```
          "items":
          {
            "type": "string"
          }
        },
        "groupFilter":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "leadFilter":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "productFilter":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "opportunityFilter":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "partnerFilter":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "pricebookFilter":{
          "type": "array",
          "items":
          {
```

```
            "type": "string"
          }
        },
        "ideaFilter":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "profileFilter":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "taskFilter":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "solutionFilter":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "userFilter":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "chatterFilter":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
```

```
        },
        "documentFilter":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "knowledgeArticleFilter":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "customEntities":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "isCrawlAccount": {
          "type": "boolean"
        },
        "isCrawlContact": {
          "type": "boolean"
        },
        "isCrawlCase": {
          "type": "boolean"
        },
        "isCrawlCampaign": {
          "type": "boolean"
        },
        "isCrawlProduct": {
          "type": "boolean"
        },
        "isCrawlLead": {
          "type": "boolean"
        },
        "isCrawlContract": {
          "type": "boolean"
        },
        "isCrawlPartner": {
```

```
          "type": "boolean"
        },
        "isCrawlProfile": {
          "type": "boolean"
        },
        "isCrawlIdea": {
          "type": "boolean"
        },
        "isCrawlPricebook": {
          "type": "boolean"
        },
        "isCrawlDocument": {
          "type": "boolean"
        },
        "crawlSharedDocument": {
          "type": "boolean"
        },
        "isCrawlGroup": {
          "type": "boolean"
        },
        "isCrawlOpportunity": {
          "type": "boolean"
        },
        "isCrawlChatter": {
          "type": "boolean"
        },
        "isCrawlUser": {
          "type": "boolean"
        },
        "isCrawlSolution":{
          "type": "boolean"
        },
        "isCrawlTask":{
          "type": "boolean"
        },

        "isCrawlAccountAttachments": {
          "type": "boolean"
        },
        "isCrawlContactAttachments": {
          "type": "boolean"
        },
        "isCrawlCaseAttachments": {
          "type": "boolean"
```

```
        },
        "isCrawlCampaignAttachments": {
          "type": "boolean"
        },
        "isCrawlLeadAttachments": {
          "type": "boolean"
        },
        "isCrawlContractAttachments": {
          "type": "boolean"
        },
        "isCrawlGroupAttachments": {
          "type": "boolean"
        },
        "isCrawlOpportunityAttachments": {
          "type": "boolean"
        },
        "isCrawlChatterAttachments": {
          "type": "boolean"
        },
        "isCrawlSolutionAttachments":{
          "type": "boolean"
        },
        "isCrawlTaskAttachments":{
          "type": "boolean"
        },
        "isCrawlCustomEntityAttachments":{
          "type": "boolean"
        },
        "isCrawlKnowledgeArticles": {
          "type": "object",
          "properties":
          {
            "isCrawlDraft": {
              "type": "boolean"
            },
            "isCrawlPublish": {
              "type": "boolean"
            },
            "isCrawlArchived": {
              "type": "boolean"
            }
          }
        },
        "inclusionDocumentFileTypePatterns":{
```

```
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "exclusionDocumentFileTypePatterns": {
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "inclusionDocumentFileNamePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "exclusionDocumentFileNamePatterns": {
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "inclusionAccountFileTypePatterns": {
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "exclusionAccountFileTypePatterns": {
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "inclusionAccountFileNamePatterns":{
          "type": "array",
          "items":
```

```
        {
          "type": "string"
        }
      },
      "exclusionAccountFileNamePatterns":{
        "type": "array",
        "items":
        {
          "type": "string"
        }
      },
      "inclusionCampaignFileTypePatterns": {
        "type": "array",
        "items":
        {
          "type": "string"
        }
      },
      "exclusionCampaignFileTypePatterns": {
        "type": "array",
        "items":
        {
          "type": "string"
        }
      },
      "inclusionCampaignFileNamePatterns":{
        "type": "array",
        "items":
        {
          "type": "string"
        }
      },
      "exclusionCampaignFileNamePatterns":{
        "type": "array",
        "items":
        {
          "type": "string"
        }
      },
      "inclusionCaseFileTypePatterns":{
        "type": "array",
        "items":
        {
          "type": "string"
```

```
        }
      },
      "exclusionCaseFileTypePatterns":{
        "type": "array",
        "items":
        {
          "type": "string"
        }
      },
      "inclusionCaseFileNamePatterns":{
        "type": "array",
        "items":
        {
          "type": "string"
        }
      },
      "exclusionCaseFileNamePatterns":{
        "type": "array",
        "items":
        {
          "type": "string"
        }
      },
      "inclusionContactFileTypePatterns":{
        "type": "array",
        "items":
        {
          "type": "string"
        }
      },
      "exclusionContactFileTypePatterns":{
        "type": "array",
        "items":
        {
          "type": "string"
        }
      },
      "inclusionContactFileNamePatterns":{
        "type": "array",
        "items":
        {
          "type": "string"
        }
      },
```

```
        "exclusionContactFileNamePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "inclusionContractFileTypePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "exclusionContractFileTypePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "inclusionContractFileNamePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "exclusionContractFileNamePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "inclusionLeadFileTypePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "exclusionLeadFileTypePatterns":{
          "type": "array",
```

```
          "items":
          {
            "type": "string"
          }
        },
        "inclusionLeadFileNamePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "exclusionLeadFileNamePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "inclusionOpportunityFileTypePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "exclusionOpportunityFileTypePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "inclusionOpportunityFileNamePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "exclusionOpportunityFileNamePatterns":{
          "type": "array",
          "items":
          {
```

```
            "type": "string"
          }
        },
        "inclusionSolutionFileTypePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "exclusionSolutionFileTypePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "inclusionSolutionFileNamePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "exclusionSolutionFileNamePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "inclusionTaskFileTypePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "exclusionTaskFileTypePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
```

```
        },
        "inclusionTaskFileNamePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "exclusionTaskFileNamePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "inclusionGroupFileTypePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "exclusionGroupFileTypePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "inclusionGroupFileNamePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "exclusionGroupFileNamePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "inclusionChatterFileTypePatterns":{
```

```
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "exclusionChatterFileTypePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "inclusionChatterFileNamePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "exclusionChatterFileNamePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "inclusionCustomEntityFileTypePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "exclusionCustomEntityFileTypePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "inclusionCustomEntityFileNamePatterns":{
          "type": "array",
          "items":
```

```
          {
            "type": "string"
          }
        },
        "exclusionCustomEntityFileNamePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        }
      },
      "required":
      []
    },
    "enableIdentityCrawler": {
      "type": "boolean"
    },
    "type": {
      "type": "string",
      "pattern": "SALESFORCEV2"
    },
    "syncMode": {
      "type": "string",
      "enum": [
        "FULL_CRAWL",
        "FORCED_FULL_CRAWL",
        "CHANGE_LOG"
      ]
    },
    "secretArn": {
      "type": "string",
      "minLength": 20,
      "maxLength": 2048
    }
  },
  "version": {
    "type": "string",
    "anyOf": [
      {
        "pattern": "1.0.0"
      }
    ]
  },
```

```
  "required": [
    "connectionConfiguration",
    "repositoryConfigurations",
    "syncMode",
    "additionalProperties",
    "secretArn",
    "type"
  ]
}
```

ServiceNow schéma de modèle

Vous incluez un JSON qui contient le schéma de la source de données dans le cadre de l'TemplateConfigurationobjet. Vous fournissez l'URL de l' ServiceNow hôte, le type d'authentification et la version de l'instance dans le cadre de la configuration de la connexion ou des détails du point de terminaison du référentiel. Spécifiez également le type de source de donnéesSERVICENOWV2, un secret pour vos informations d'authentification et les autres configurations nécessaires. Vous spécifiez ensuite TEMPLATE le Type moment de l'appel CreateDataSource.

Vous pouvez utiliser le modèle fourni dans ce guide du développeur. Consultez ServiceNow Schéma JSON.

Le tableau suivant décrit les paramètres du schéma ServiceNow JSON.

Configuration	Description
Configuration de connexion	Informations de configuration pour le point de terminaison de la source de données.
repositoryEndpointMetadata	Informations sur le point de terminaison de la source de données.
URL de l'hôte	URL de l' ServiceNow hôte. Par exemple, *your-domain.service-now.com* .
authType	Le type d'authentification que vous utilisez, que ce soit basicAuth ouOAuth2.
servicenowInstanceVersion	La ServiceNow version que vous utilisez. Vous pouvez choisir entre Tokyo Sandiego,Rome, etOthers.

Configuration	Description
Configurations du référentiel	Informations de configuration pour le contenu de la source de données. Par exemple, configurer des types spécifiques de contenu et des mappages de champs.
• Article de connaissances • attachment • Catalogue de services • incident	Liste d'objets qui mappent les attributs ou les noms de champs de vos articles de ServiceNow connaissa nces, de vos pièces jointes, de votre catalogue de services et de vos incidents pour Amazon Kendra in dexer les noms de champs. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données. Les noms des champs de source de ServiceNow données doivent figurer dans vos métadonnées ServiceNow personnalisées.
propriétés supplémentaires	Options de configuration supplémentaires pour votre contenu dans votre source de données.
maxFileSizeInMegaBytes	Spécifiez la limite de taille de fichier MBs qu'Amazon Kendra explorera. Amazon Kendra explorera uniquemen t les fichiers dans la limite de taille que vous avez définie. La taille de fichier par défaut est de 50 Mo. La taille maximale du fichier doit être supérieure à 0 Mo et inférieure ou égale à 50 Mo.

Configuration	Description
• knowledgeArticleFilter • incidentQueryFilter • serviceCatalogQueryFiltre • knowledgeArticleTitleRegExp • serviceCatalogTitleRegExp • incidentTitleRegExp • inclusionFileTypeMotifs • exclusionFileTypeMotifs • inclusionFileNameMotifs • exclusionFileNameMotifs • incidentStateType	Liste de modèles d'expressions régulières à inclure et à and/or exclure certains fichiers de votre source de ServiceNow données. Les fichiers qui correspondent aux modèles sont inclus dans l'index. Les fichiers qui ne correspondent pas aux modèles sont exclus de l'index. Si un fichier correspond à la fois à un modèle d'inclusi on et d'exclusion, le modèle d'exclusion a la priorité et le fichier n'est pas inclus dans l'index.
• isCrawlKnowledgeL'article • isCrawlKnowledgeArticleAttachment • includePublicArticlesUniquement • isCrawlServiceCatalogue • isCrawlServiceCatalogAttachment • isCrawlActiveServiceCatalog • isCrawlInactiveServiceCatalog • isCrawlIncident • isCrawlIncidentPièce jointe • isCrawlActiveIncident • isCrawlInactiveIncident • appliquer ACLFor KnowledgeArticle • appliquer ACLFor ServiceCatalog • appliquer ACLFor Incident	`true`pour parcourir les articles de ServiceNow connaissances, les catalogues de services, les incidents et les pièces jointes.
type	Type de source de données. Spécifiez `SERVICENOWV2` comme type de source de données.

Configuration	Description
enableIdentityCrawler	`true`utiliser Amazon Kendra le moteur de recherche d'identité pour synchroniser les identity/principal informations relatives aux utilisateurs et aux groupes ayant accès à certains documents. Si le robot d'identif ication est désactivé, tous les documents peuvent être consultés publiquement. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'identification est désactivé, vous pouvez également utiliser l'PutPrincipalMappingAPI pour télécharger les informations d'accès des utilisateurs et des groupes.
Mode de synchronisation	Spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Vous pouvez choisir entre : • `FORCED_FULL_CRAWL` pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index. • `FULL_CRAWL` pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

Configuration	Description

Secrétaire Arn	Le nom de ressource Amazon (ARN) d'un AWS Secrets Manager secret contenant les paires clé-valeur requises pour se connecter à votre. ServiceNow Le secret doit contenir une structure JSON avec les clés suivantes :
	`{ "username": " user name", "password": " password" }`
	Si vous utilisez OAuth2 l'authentification, votre secret doit contenir une structure JSON avec les clés suivantes :
	`{ "username": " user name", "password": " password", "clientId": " client id", "clientSecret": " client secret" }`

version	Version du modèle actuellement prise en charge.

ServiceNow Schéma JSON

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "hostUrl": {
              "type": "string",
```

```
              "pattern": "^(?!(^(https?|ftp|file):\/\/))[a-z0-9-]+(.service-
now.com|.servicenowservices.com)$",
              "minLength": 1,
              "maxLength": 2048
            },
            "authType": {
              "type": "string",
              "enum": [
                "basicAuth",
                "OAuth2"
              ]
            },
            "servicenowInstanceVersion": {
              "type": "string",
              "enum": [
                "Tokyo",
                "Sandiego",
                "Rome",
                "Others"
                ]
            }
          },
          "required": [
            "hostUrl",
            "authType",
            "servicenowInstanceVersion"
          ]
        }
      },
      "required": [
        "repositoryEndpointMetadata"
      ]
    },
    "repositoryConfigurations": {
      "type": "object",
      "properties": {
        "knowledgeArticle": {
          "type": "object",
          "properties": {
            "fieldMappings": {
              "type": "array",
              "items": [
                {
                  "type": "object",
```

```
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string",
                    "enum": [
                      "STRING",
                      "DATE",
                      "STRING_LIST"
                    ]
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  },
                  "dateFieldFormat": {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required": [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      },
      "attachment": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
```

```
                },
                "indexFieldType": {
                  "type": "string",
                  "enum": [
                    "STRING",
                    "LONG",
                    "DATE",
                    "STRING_LIST"
                  ]
                },
                "dataSourceFieldName": {
                  "type": "string"
                },
                "dateFieldFormat": {
                  "type": "string",
                  "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                }
              },
              "required": [
                "indexFieldName",
                "indexFieldType",
                "dataSourceFieldName"
              ]
            }
          ]
        }
      },
      "required": [
        "fieldMappings"
      ]
    },
    "serviceCatalog": {
      "type": "object",
      "properties": {
        "fieldMappings": {
          "type": "array",
          "items": [
            {
              "type": "object",
              "properties": {
                "indexFieldName": {
                  "type": "string"
                },
                "indexFieldType": {
```

```
                    "type": "string",
                    "enum": [
                      "STRING",
                      "DATE",
                      "STRING_LIST"
                    ]
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  },
                  "dateFieldFormat": {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required": [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      },
      "incident": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string",
                    "enum": [
                      "STRING",
```

```
                    "DATE",
                    "STRING_LIST"
                  ]
                },
                "dataSourceFieldName": {
                  "type": "string"
                },
                "dateFieldFormat": {
                  "type": "string",
                  "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                }
              },
              "required": [
                "indexFieldName",
                "indexFieldType",
                "dataSourceFieldName"
              ]
            }
          ]
        }
      },
      "required": [
        "fieldMappings"
      ]
    }
  }
},
"additionalProperties": {
  "type": "object",
  "properties": {
    "maxFileSizeInMegaBytes": {
      "type": "string"
    },
    "isCrawlKnowledgeArticle": {
      "type": "boolean"
    },
    "isCrawlKnowledgeArticleAttachment": {
      "type": "boolean"
    },
    "includePublicArticlesOnly": {
      "type": "boolean"
    },
    "knowledgeArticleFilter": {
      "type": "string"
```

```
        },
        "incidentQueryFilter": {
          "type": "string"
        },
        "serviceCatalogQueryFilter": {
          "type": "string"
        },
        "isCrawlServiceCatalog": {
          "type": "boolean"
        },
        "isCrawlServiceCatalogAttachment": {
          "type": "boolean"
        },
        "isCrawlActiveServiceCatalog": {
          "type": "boolean"
        },
        "isCrawlInactiveServiceCatalog": {
          "type": "boolean"
        },
        "isCrawlIncident": {
          "type": "boolean"
        },
        "isCrawlIncidentAttachment": {
          "type": "boolean"
        },
        "isCrawlActiveIncident": {
          "type": "boolean"
        },
        "isCrawlInactiveIncident": {
          "type": "boolean"
        },
        "applyACLForKnowledgeArticle": {
          "type": "boolean"
        },
        "applyACLForServiceCatalog": {
          "type": "boolean"
        },
        "applyACLForIncident": {
          "type": "boolean"
        },
        "incidentStateType": {
          "type": "array",
          "items": {
            "type": "string",
```

```
            "enum": [
              "Open",
              "Open - Unassigned",
              "Resolved",
              "All"
            ]
          }
        },
        "knowledgeArticleTitleRegExp": {
          "type": "string"
        },
        "serviceCatalogTitleRegExp": {
          "type": "string"
        },
        "incidentTitleRegExp": {
          "type": "string"
        },
        "inclusionFileTypePatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "exclusionFileTypePatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "inclusionFileNamePatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "exclusionFileNamePatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        }
      },
      "required": []
    },
```

```
    "type": {
      "type": "string",
      "pattern": "SERVICENOWV2"
    },
    "enableIdentityCrawler": {
      "type": "boolean"
    },
    "syncMode": {
      "type": "string",
      "enum": [
        "FORCED_FULL_CRAWL",
        "FULL_CRAWL"
      ]
    },
    "secretArn": {
      "type": "string",
      "minLength": 20,
      "maxLength": 2048
    }
  },
  "version": {
    "type": "string",
    "anyOf": [
      {
        "pattern": "1.0.0"
      }
    ]
  },
  "required": [
    "connectionConfiguration",
    "repositoryConfigurations",
    "syncMode",
    "additionalProperties",
    "secretArn",
    "type"
  ]
}
```

Schéma du modèle Slack

Vous incluez un JSON qui contient le schéma de la source de données dans le cadre de TemplateConfigurationl'objet. Vous fournissez l'URL de l'hôte dans le cadre de la configuration de la connexion ou des détails du point de terminaison du référentiel. Spécifiez également le

type de source de donnéesSLACK, un secret pour vos informations d'authentification et les autres configurations nécessaires. Vous spécifiez ensuite TEMPLATE le Type moment de l'appel CreateDataSource.

Vous pouvez utiliser le modèle fourni dans ce guide du développeur. Consultez Schéma JSON Slack.

Le tableau suivant décrit les paramètres du schéma JSON de Slack.

Configuration	Description
Configuration de connexion	Informations de configuration pour le point de terminaison de la source de données.
repositoryEndpointMetadata	Informations sur le point de terminaison de la source de données.
Identifiant de l'équipe	L'identifiant d'équipe Slack que vous avez copié depuis l'URL de votre page principale de Slack.
Configurations du référentiel	Informations de configuration pour le contenu de la source de données. Par exemple, configurer des types spécifiques de contenu et des mappages de champs.
Tous	Liste d'objets qui mappent les attributs ou les noms de champs de votre Slack contenu aux noms de champs d' Amazon Kendra index.
Propriétés supplémentaires	Options de configuration supplémentaires pour votre contenu dans votre source de données.
Modèles d'inclusion	Liste de modèles d'expressions régulières permettant d'inclure un contenu spécifique dans votre source de Slack données. Le contenu qui correspond aux modèles est inclus dans l'index. Le contenu qui ne correspond pas aux modèles est exclu de l'index. Si un contenu correspond à la fois à un modèle d'inclusion et à un modèle

Configuration	Description
	d'exclusion, le modèle d'exclusion est prioritaire et le contenu n'est pas inclus dans l'index.
Schémas d'exclusion	Liste de modèles d'expressions régulières permettant d'exclure un contenu spécifique de votre source de Slack données. Le contenu qui correspond aux modèles est exclu de l'index. Le contenu qui ne correspond pas aux modèles est inclus dans l'index. Si un contenu correspon d à la fois à un modèle d'inclusion et à un modèle d'exclusion, le modèle d'exclusion est prioritaire et le contenu n'est pas inclus dans l'index.
crawlBotMessages	`true`pour explorer les messages du bot.
Exclure Archivé	`true`pour exclure l'exploration des messages archivés.
Type de conversation	Le type de conversation que vous souhaitez indexer si `PUBLIC_CHANNEL PRIVATE_C HANNEL` , `GROUP_MESSAGE` et`DIRECT_ME SSAGE` .
Filtre de canal	Type de canal que vous souhaitez indexer `private_channel` ou`public_channel` .
Depuis Date	Vous pouvez choisir de configurer un `sinceDate` paramètre afin que le Slack connecteur analyse le contenu en fonction d'un paramètre spécifique`sinceDate` .

Configuration	Description
Regardez en arrière	Vous pouvez choisir de configurer un `lookBack` paramètre afin que le Slack connecteur analyse le contenu mis à jour ou supprimé jusqu'à un certain nombre d'heures avant la dernière synchronisation du connecteu r.
Mode de synchronisation	Spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Vous pouvez choisir entre : • `FORCED_FULL_CRAWL` pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index. • `FULL_CRAWL` pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modificat ions de contenu et indexer le contenu modifié depuis la dernière synchronisation. • `CHANGE_LOG` pour indexer uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchroni se avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

Configuration	Description
type	Type de source de données. Spécifiez `SLACK` comme type de source de données.
enableIdentityCrawler	`true`utiliser Amazon Kendra le moteur de recherche d'identité pour synchroniser les identity/principal informations relatives aux utilisateurs et aux groupes ayant accès à certains documents. Si le robot d'identification est désactivé, tous les documents peuvent être consultés publiquement. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'identification est désactivé , vous pouvez également utiliser l'PutPrincipalMappingAPI pour télécharger les informations d'accès des utilisateurs et des groupes.
Secrétaire Arn	Le nom de ressource Amazon (ARN) d'un AWS Secrets Manager secret contenant les paires clé-valeur requises pour se connecter à votre. Slack Le secret doit contenir une structure JSON avec les clés suivantes : `{ "slackToken": " token" }`
version	Version de ce modèle actuellement prise en charge.

Schéma JSON Slack

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
```

```
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "teamId": {
              "type": "string"
            }
          },
          "required": ["teamId"]
        }
      }
    },
    "repositoryConfigurations": {
      "type": "object",
      "properties": {
        "All": {
          "type": "object",
          "properties": {
            "fieldMappings": {
              "type": "array",
              "items": [
                {
                  "type": "object",
                  "properties": {
                    "indexFieldName": {
                      "type": "string"
                    },
                    "indexFieldType": {
                      "type": "string",
                      "enum": ["STRING", "STRING_LIST", "DATE","LONG"]
                    },
                    "dataSourceFieldName": {
                      "type": "string"
                    },
                    "dateFieldFormat": {
                      "type": "string",
                      "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                    }
                  },
                  "required": [
                    "indexFieldName",
                    "indexFieldType",
                    "dataSourceFieldName"
```

```
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      }
    },
    "required": [
    ]
  },
  "additionalProperties": {
    "type": "object",
    "properties": {
      "exclusionPatterns": {
        "type": "array",
        "items": {
          "type": "string"
        }
      },
      "inclusionPatterns": {
        "type": "array",
        "items": {
          "type": "string"
        }
      },
      "crawlBotMessages": {
        "type": "boolean"
      },
      "excludeArchived": {
        "type": "boolean"
      },
      "conversationType": {
        "type": "array",
        "items": {
          "type": "string",
          "enum": [
            "PUBLIC_CHANNEL",
            "PRIVATE_CHANNEL",
            "GROUP_MESSAGE",
            "DIRECT_MESSAGE"
          ]
```

```
          }
        },
        "channelFilter": {
            "type": "object",
            "properties": {
              "private_channel": {
                "type": "array",
                "items": {
                  "type": "string"
                }
              },
              "public_channel": {
                "type": "array",
                "items": {
                  "type": "string"
                }
              }
          }
        },
        "channelIdFilter": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "sinceDate": {
          "anyOf": [
            {
              "type": "string",
              "pattern": "^[0-9]{4}-[0-9]{2}-[0-9]{2}T[0-9]{2}:[0-9]{2}:[0-9]{2}Z$"
            },
            {
              "type": "string",
              "pattern": ""
            }
          ]
        },
        "lookBack": {
          "type": "string",
          "pattern": "^[0-9]*$"
        }
      },
      "required": [
      ]
```

```
    },
    "syncMode": {
      "type": "string",
      "enum": [
        "FORCED_FULL_CRAWL",
        "FULL_CRAWL",
        "CHANGE_LOG"
      ]
    },
    "type" : {
      "type" : "string",
      "pattern": "SLACK"
    },
    "enableIdentityCrawler": {
      "type": "boolean"
    },
    "secretArn": {
      "type": "string"
    }
  },
  "version": {
    "type": "string",
    "anyOf": [
      {
        "pattern": "1.0.0"
      }
    ]
  },
  "required": [
    "connectionConfiguration",
    "repositoryConfigurations",
    "syncMode",
    "additionalProperties",
    "secretArn",
    "type",
    "enableIdentityCrawler"
  ]
}
```

Schéma du modèle Zendesk

Vous incluez un JSON qui contient le schéma de la source de données dans le cadre de TemplateConfigurationl'objet. Vous fournissez l'URL de l'hôte dans le cadre de la configuration

de la connexion ou des détails du point de terminaison du référentiel. Spécifiez également le type de source de donnéesZENDESK, un secret pour vos informations d'authentification et les autres configurations nécessaires. Vous spécifiez ensuite TEMPLATE le Type moment de l'appel CreateDataSource.

Vous pouvez utiliser le modèle fourni dans ce guide du développeur. Consultez Schéma JSON Zendesk.

Le tableau suivant décrit les paramètres du schéma Zendesk JSON.

Configuration	Description
Configuration de connexion	Informations de configuration pour le point de terminaison de la source de données.
repositoryEndpointMetadata	Informations sur le point de terminaison de la source de données.
URL de l'hôte	URL de l'hôte Zendesk. Par exemple, https://yoursubdomain.zendesk.com.
Configurations du référentiel	Informations de configuration pour le contenu de la source de données. Par exemple, configurer des types spécifiques de contenu et des mappages de champs.
• ticket • Commentaire sur le billet • ticketCommentAttachment • article • Commentaire de l'article • Pièce jointe à l'article • Thème communautaire • communityPostComment	Liste d'objets qui associent les attributs ou les noms de champs des tickets Zendesk aux noms des champs d'index Amazon Kendra. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.
Secrétaire Arn	Le nom de ressource Amazon (ARN) d'un AWS Secrets Manager secret qui contient les paires clé-valeur requises pour se connecter à votre

Configuration	Description
	Zendesk. Le secret doit contenir une structure JSON avec les clés suivantes : URL hôte, ID client, secret client, nom d'utilisateur et mot de passe.
Propriétés supplémentaires	Options de configuration supplémentaires pour votre contenu dans votre source de données
organizationNameFilter	Vous pouvez choisir d'indexer les tickets qui existent au sein d'une organisation spécifique.
Depuis Date	Vous pouvez choisir de configurer un `sinceDate` paramètre afin que le connecteur Zendesk analyse le contenu en fonction d'un paramètre spécifique. `sinceDate`
Modèles d'inclusion	Une liste de modèles d'expressions régulières pour inclure certains fichiers dans votre source de données Zendesk. Les fichiers qui correspondent aux modèles sont inclus dans l'index. Les fichiers qui ne correspondent pas aux modèles sont exclus de l'index. Si un fichier correspond à la fois à un modèle d'inclusion et à un modèle d'exclusion, le modèle d'exclusion est prioritaire et le fichier n'est pas inclus dans l'index.

Configuration	Description
Schémas d'exclusion	Une liste de modèles d'expressions régulière s pour exclure certains fichiers de votre source de données Zendesk. Les fichiers qui correspondent aux modèles sont exclus de l'index. Les fichiers qui ne correspondent pas aux modèles sont inclus dans l'index. Si un fichier correspond à la fois à un modèle d'exclusion et à un modèle d'inclusion, le modèle d'exclusion est prioritaire et le fichier n'est pas inclus dans l'index.
• isCrawlTicket • isCrawlTicketCommentaire • isCrawlTicketCommentAttachment • isCrawlArticle • isCrawlArticleCommentaire • isCrawlArticlePièce jointe • isCrawlCommunityRubrique • isCrawlCommunityPublier • isCrawlCommunityPostComment	Entrez « `true` » pour explorer ces types de contenu.
type	Spécifiez `ZENDESK` comme type de source de données.
useChangeLog	Entrez « `true` » pour utiliser le journal des modifications de Zendesk afin de déterminer quels documents doivent être mis à jour dans l'index. Selon la taille du journal des modificat ions, il peut être plus rapide de numériser les documents dans Zendesk. Si vous synchroni sez votre source de données Zendesk avec votre index pour la première fois, tous les documents sont numérisés.

Schéma JSON Zendesk

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "hostUrl": {
              "type": "string",
              "pattern": "https:.*"
            }
          },
          "required": [
            "hostUrl"
          ]
        }
      },
      "required": [
        "repositoryEndpointMetadata"
      ]
    },
    "repositoryConfigurations": {
      "type": "object",
      "properties": {
        "ticket": {
          "type": "object",
          "properties": {
            "fieldMappings": {
              "type": "array",
              "items": {
                "anyOf": [
                  {
                    "type": "object",
                    "properties": {
                      "indexFieldName": {
                        "type": "string"
                      },
                      "indexFieldType": {
                        "type": "string",
```

```
                    "enum": ["STRING", "STRING_LIST", "LONG", "DATE"]
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  },
                  "dateFieldFormat": {
                    "type": "string",
                    "pattern": "dd-MM-yyyy HH:mm:ss"
                  }
                },
                "required": [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"

                ]
              }
            ]
          }
        }
      },
      "required": [
        "fieldMappings"
      ]
    },
    "ticketComment": {
      "type": "object",
      "properties": {
        "fieldMappings": {
          "type": "array",
          "items": {
            "anyOf": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string",
                    "enum": ["STRING", "STRING_LIST", "LONG", "DATE"]
                  },
                  "dataSourceFieldName": {
                    "type": "string"
```

```
                    },
                    "dateFieldFormat": {
                      "type": "string",
                      "pattern": "dd-MM-yyyy HH:mm:ss"
                    }
                  },
                  "required": [
                    "indexFieldName",
                    "indexFieldType",
                    "dataSourceFieldName"

                  ]
                }
              ]
            }
          }
        },
        "required": [
          "fieldMappings"
        ]
      },
      "ticketCommentAttachment": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": {
              "anyOf": [
                {
                  "type": "object",
                  "properties": {
                    "indexFieldName": {
                      "type": "string"
                    },
                    "indexFieldType": {
                      "type": "string",
                      "enum": ["STRING", "STRING_LIST", "LONG", "DATE"]
                    },
                    "dataSourceFieldName": {
                      "type": "string"
                    },
                    "dateFieldFormat": {
                      "type": "string",
                      "pattern": "dd-MM-yyyy HH:mm:ss"
```

```
                    }
                  },
                  "required": [
                    "indexFieldName",
                    "indexFieldType",
                    "dataSourceFieldName"
                  ]
                }
              ]
            }
          }
        },
        "required": [
          "fieldMappings"
        ]
      },
      "article": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": {
              "anyOf": [
                {
                  "type": "object",
                  "properties": {
                    "indexFieldName": {
                      "type": "string"
                    },
                    "indexFieldType": {
                      "type": "string",
                      "enum": ["STRING", "STRING_LIST", "LONG", "DATE"]
                    },
                    "dataSourceFieldName": {
                      "type": "string"
                    },
                    "dateFieldFormat": {
                      "type": "string",
                      "pattern": "dd-MM-yyyy HH:mm:ss"
                    }
                  },
                  "required": [
                    "indexFieldName",
                    "indexFieldType",
```

```
                    "dataSourceFieldName"
                  ]
                }
              ]
            }
          }
        },
        "required": [
          "fieldMappings"
        ]
      },
      "communityPostComment": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": {
              "anyOf": [
                {
                  "type": "object",
                  "properties": {
                    "indexFieldName": {
                      "type": "string"
                    },
                    "indexFieldType": {
                      "type": "string",
                      "enum": ["STRING", "STRING_LIST", "LONG", "DATE"]
                    },
                    "dataSourceFieldName": {
                      "type": "string"
                    },
                    "dateFieldFormat": {
                      "type": "string",
                      "pattern": "dd-MM-yyyy HH:mm:ss"
                    }
                  },
                  "required": [
                    "indexFieldName",
                    "indexFieldType",
                    "dataSourceFieldName"
                  ]
                }
              ]
            }
```

```
            }
          },
          "required": [
            "fieldMappings"
          ]
        },
        "articleComment": {
          "type": "object",
          "properties": {
            "fieldMappings": {
              "type": "array",
              "items": {
                "anyOf": [
                  {
                    "type": "object",
                    "properties": {
                      "indexFieldName": {
                        "type": "string"
                      },
                      "indexFieldType": {
                        "type": "string",
                        "enum": ["STRING", "STRING_LIST", "LONG", "DATE"]
                      },
                      "dataSourceFieldName": {
                        "type": "string"
                      },
                      "dateFieldFormat": {
                        "type": "string",
                        "pattern": "dd-MM-yyyy HH:mm:ss"
                      }
                    },
                    "required": [
                      "indexFieldName",
                      "indexFieldType",
                      "dataSourceFieldName"
                    ]
                  }
                ]
              }
            }
          },
          "required": [
            "fieldMappings"
          ]
```

```
        },
        "articleAttachment": {
          "type": "object",
          "properties": {
            "fieldMappings": {
              "type": "array",
              "items": {
                "anyOf": [
                  {
                    "type": "object",
                    "properties": {
                      "indexFieldName": {
                        "type": "string"
                      },
                      "indexFieldType": {
                        "type": "string",
                        "enum": ["STRING", "STRING_LIST", "LONG", "DATE"]
                      },
                      "dataSourceFieldName": {
                        "type": "string"
                      },
                      "dateFieldFormat": {
                        "type": "string",
                        "pattern": "dd-MM-yyyy HH:mm:ss"
                      }
                    },
                    "required": [
                      "indexFieldName",
                      "indexFieldType",
                      "dataSourceFieldName"
                    ]
                  }
                ]
              }
            }
          },
          "required": [
            "fieldMappings"
          ]
        },
        "communityTopic": {
          "type": "object",
          "properties": {
            "fieldMappings": {
```

```
            "type": "array",
            "items": {
              "anyOf": [
                {
                  "type": "object",
                  "properties": {
                    "indexFieldName": {
                      "type": "string"
                    },
                    "indexFieldType": {
                      "type": "string",
                      "enum": ["STRING", "STRING_LIST", "LONG", "DATE"]
                    },
                    "dataSourceFieldName": {
                      "type": "string"
                    },
                    "dateFieldFormat": {
                      "type": "string",
                      "pattern": "dd-MM-yyyy HH:mm:ss"
                    }
                  },
                  "required": [
                    "indexFieldName",
                    "indexFieldType",
                    "dataSourceFieldName"
                  ]
                }
              ]
            }
          }
        },
        "required": [
          "fieldMappings"
        ]
      }
    }
  },
  "secretArn": {
    "type": "string",
    "minLength": 20,
    "maxLength": 2048
  },
  "additionalProperties": {
    "type": "object",
```

```
      "properties": {
        "organizationNameFilter": {
          "type": "array"
        },
        "sinceDate": {
          "type": "string",
          "pattern": "^[0-9]{4}-[0-9]{2}-[0-9]{2} [0-9]{2}:[0-9]{2}:[0-9]{2}$"
        },
        "inclusionPatterns": {
          "type": "array"
        },
        "exclusionPatterns": {
          "type": "array"
        },
        "isCrawTicket": {
          "type": "string"
        },
        "isCrawTicketComment": {
          "type": "string"
        },
        "isCrawTicketCommentAttachment": {
          "type": "string"
        },
        "isCrawlArticle": {
          "type": "string"
        },
        "isCrawlArticleAttachment": {
          "type": "string"
        },
        "isCrawlArticleComment": {
          "type": "string"
        },
        "isCrawlCommunityTopic": {
          "type": "string"
        },
        "isCrawlCommunityPost": {
          "type": "string"
        },
        "isCrawlCommunityPostComment": {
          "type": "string"
        }
      }
    },
    "type": {
```

```
      "type": "string",
      "pattern": "ZENDESK"
    },
    "useChangeLog": {
      "type": "string",
      "enum": ["true", "false"]
    }
  },
  "version": {
    "type": "string",
    "anyOf": [
      {
        "pattern": "1.0.0"
      }
    ]
  },
  "additionalProperties": false,
  "required": [
    "connectionConfiguration",
    "repositoryConfigurations",
    "additionalProperties",
    "useChangeLog",
    "secretArn",
    "type"
  ]
}
```

Adobe Experience Manager

> **Note**
>
> Le connecteur Adobe Experience Manager reste entièrement pris en charge pour les clients existants jusqu'au 31 mai 2026. Bien que ce connecteur ne soit plus disponible pour les nouveaux utilisateurs, les utilisateurs actuels peuvent continuer à l'utiliser sans interruption. Nous faisons constamment évoluer notre portefeuille de connecteurs afin de proposer des solutions plus évolutives et personnalisables. Pour les intégrations futures, nous vous recommandons d'explorer le framework de connecteurs personnalisés Amazon Kendra [1], conçu pour prendre en charge un plus large éventail de cas d'utilisation en entreprise avec une flexibilité accrue.

Adobe Experience Managerest un système de gestion de contenu utilisé pour créer du contenu de sites Web ou d'applications mobiles. Vous pouvez l'utiliser Amazon Kendra pour vous connecter à vos pages Adobe Experience Manager et à vos actifs de contenu et les indexer.

Amazon Kendra prend en charge Adobe Experience Manager (AEM) en tant qu'instance d'auteur du service Adobe Experience Manager cloud et en tant qu'instance de création et de publication sur site.

Vous pouvez vous connecter Amazon Kendra à votre source de Adobe Experience Manager données à l'aide de la Amazon Kendra console ou de l'TemplateConfigurationAPI.

Pour résoudre les problèmes liés à votre connecteur de source de données Amazon Kendra Adobe Experience Manager, consultezDépannage des sources de données.

Rubriques

- Fonctionnalités prises en charge
- Prérequis
- Instructions de connexion

Fonctionnalités prises en charge

Adobe Experience Managerle connecteur de source de données prend en charge les fonctionnalités suivantes :

- Mappages de champs
- Contrôle d'accès utilisateur
- Filtres d'inclusion/exclusion
- Synchronisation complète et incrémentielle du contenu
- OAuth Authentification 2.0 et de base
- Cloud privé virtuel (VPC)

Prérequis

Avant de pouvoir Amazon Kendra indexer votre source de Adobe Experience Manager données, apportez ces modifications à vos AWS comptes Adobe Experience Manager and.

DansAdobe Experience Manager, assurez-vous d'avoir :

- Accès à un compte doté de privilèges administratifs ou à un utilisateur administrateur.
- Vous avez copié Adobe Experience Manager l'URL de votre hôte.

> **Note**
>
> (Sur place/sur serveur) Amazon Kendra vérifie si les informations de point de terminaison incluses sont les mêmes AWS Secrets Manager que celles spécifiées dans les détails de configuration de votre source de données. Cela permet de se protéger contre le problème de confusion des adjoints, qui est un problème de sécurité lorsqu'un utilisateur n'est pas autorisé à effectuer une action mais l'utilise Amazon Kendra comme proxy pour accéder au secret configuré et exécuter l'action. Si vous modifiez ultérieurement les informations de votre point de terminaison, vous devez créer un nouveau secret pour synchroniser ces informations.

- Vous avez noté vos informations d'authentification de base, à savoir le nom d'utilisateur et le mot de passe de l'administrateur.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

- Facultatif : informations d'identification OAuth 2.0 configurées dans Adobe Experience Manager (AEM) en tant que service cloud ou AEM sur site. Si vous utilisez AEM On-Premise, les informations d'identification incluent l'ID client, le secret client et la clé privée. Si vous utilisez AEM en tant que service cloud, les informations d'identification incluent l'identifiant du client, le secret du client, la clé privée, l'identifiant de l'organisation, l'identifiant du compte technique et l'hôte Adobe Identity Management System (IMS). Pour plus d'informations sur la façon de générer ces informations d'identification pour AEM as a Cloud Service, consultez Adobe Experience Managerla documentation. Pour AEM On-Premise, l'implémentation du serveur Adobe Granite OAuth 2.0 (com.adobe.granite.oauth.server) prend en charge les fonctionnalités du serveur OAuth 2.0 dans AEM.
- Il est vérifié que chaque document est unique dans Adobe Experience Manager et dans les autres sources de données que vous prévoyez d'utiliser pour le même index. Chaque source de données

que vous souhaitez utiliser pour un index ne doit pas contenir le même document dans toutes les sources de données. IDs Les documents sont globaux par rapport à un index et doivent être uniques par index.

Dans votre Compte AWS, assurez-vous d'avoir :

- Création d'un Amazon Kendra index et, si vous utilisez l'API, notez l'ID de l'index.
- Vous avez créé un IAM rôle pour votre source de données et, si vous utilisez l'API, notez l'ARN du IAM rôle.

> **Note**
>
> Si vous modifiez votre type d'authentification et vos informations d'identification, vous devez mettre à jour votre IAM rôle pour accéder au bon identifiant AWS Secrets Manager secret.

- Stockez vos informations d'authentification Adobe Experience Manager dans un AWS Secrets Manager secret et, si vous utilisez l'API, notez l'ARN du secret.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

Si vous n'avez pas de IAM rôle ou de secret existant, vous pouvez utiliser la console pour créer un nouveau IAM rôle et un nouveau Secrets Manager secret lorsque vous connectez votre source de données Adobe Experience Manager à Amazon Kendra. Si vous utilisez l'API, vous devez fournir l'ARN d'un IAM rôle et d'un Secrets Manager secret existants, ainsi qu'un identifiant d'index.

Instructions de connexion

Pour vous connecter Amazon Kendra à votre source de Adobe Experience Manager données, vous devez fournir les informations nécessaires sur votre source de Adobe Experience Manager données

afin de Amazon Kendra pouvoir accéder à vos données. Si vous n'avez pas encore configuré Adobe Experience Manager pour Amazon Kendra, consultezPrérequis.

Console

Pour vous connecter Amazon Kendra à Adobe Experience Manager

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la.
2. Dans le volet de navigation de gauche, choisissez Index, puis choisissez l'index que vous souhaitez utiliser dans la liste des index.

> **Note**
>
> Vous pouvez choisir de configurer ou de modifier vos paramètres de contrôle d'accès utilisateur dans les paramètres de l'index.

3. Sur la page de démarrage, choisissez Ajouter une source de données.
4. Sur la page Ajouter une source de données, choisissez le connecteur Adobe Experience Manager, puis sélectionnez Ajouter un connecteur. Si vous utilisez la version 2 (le cas échéant), choisissez le connecteur Adobe Experience Manager avec la balise « V2.0 ».
5. Sur la page Spécifier les détails de la source de données, entrez les informations suivantes :
   a. Dans Nom et description, pour Nom de la source de données : entrez le nom de votre source de données. Vous pouvez inclure des traits d'union, mais pas des espaces.
   b. (Facultatif) Description : entrez une description facultative pour votre source de données.
   c. Dans la langue par défaut : choisissez une langue pour filtrer vos documents pour l'index. Sauf indication contraire, la langue par défaut est l'anglais. La langue spécifiée dans les métadonnées du document remplace la langue sélectionnée.
   d. Dans Balises, pour Ajouter une nouvelle balise : incluez des balises facultatives pour rechercher et filtrer vos ressources ou suivre vos AWS coûts.
   e. Choisissez Suivant.
6. Sur la page Définir l'accès et la sécurité, entrez les informations suivantes :
   a. Source : choisissez AEM sur site ou AEM en tant que service cloud.

      Entrez l'URL de votre Adobe Experience Manager hôte. Par exemple, si vous utilisez AEM On-Premise, vous devez inclure le nom d'hôte et le port :. https://hostname:port Ou,

si vous utilisez AEM en tant que service cloud, vous pouvez utiliser l'URL de l'auteur : https://author-xxxxxx-xxxxxxx.adobeaemcloud.com.

b. Emplacement du certificat SSL : entrez le chemin d'accès au certificat SSL stocké dans un Amazon S3 compartiment. Vous l'utilisez pour vous connecter à AEM On-Premise avec une connexion SSL sécurisée.

c. Autorisation : activez ou désactivez les informations de la liste de contrôle d'accès (ACL) pour vos documents, si vous disposez d'une ACL et que vous souhaitez l'utiliser pour le contrôle d'accès. L'ACL indique les documents auxquels les utilisateurs et les groupes peuvent accéder. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.

d. Authentification —Choisissez l'authentification de base ou l'authentification OAuth 2.0. Choisissez ensuite un AWS Secrets Manager secret existant ou créez-en un nouveau pour stocker vos Adobe Experience Manager informations d'identification. Si vous choisissez de créer un nouveau secret, une fenêtre AWS Secrets Manager secrète s'ouvre.

   Si vous avez choisi l'authentification de base, entrez le nom du secret, le nom d'utilisateur et le mot de passe du Adobe Experience Manager site. L'utilisateur doit disposer d'une autorisation d'administrateur ou être un utilisateur administrateur.

   Si vous avez choisi l'authentification OAuth 2.0 et que vous utilisez AEM On-Premise, entrez un nom pour le secret, l'ID client, le secret client et la clé privée. Si vous utilisez AEM en tant que service cloud, entrez un nom pour le secret, l'identifiant client, le secret client, la clé privée, l'identifiant de l'organisation, l'identifiant du compte technique et l'hôte Adobe Identity Management System (IMS).

   Enregistrez et ajoutez votre secret.

e. Virtual Private Cloud (VPC) —Vous pouvez choisir d'utiliser un VPC. Dans ce cas, vous devez ajouter des sous-réseaux et des groupes de sécurité VPC.

f. Identity Crawler : spécifiez s'il faut activer l'explorateur Amazon Kendra d'identité. Le robot d'exploration d'identité utilise les informations de la liste de contrôle d'accès (ACL) de vos documents pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Si vous disposez d'une ACL pour vos documents et que vous choisissez de l'utiliser, vous pouvez également choisir d'activer le robot d'exploration Amazon Kendra d'identité pour configurer le filtrage des résultats

de recherche par contexte utilisateur. Sinon, si le robot d'identification est désactivé, tous les documents peuvent être consultés publiquement. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'exploration d'identité est désactivé, vous pouvez également utiliser l'PutPrincipalMappingAPI pour télécharger les informations d'accès des utilisateurs et des groupes afin de filtrer le contexte utilisateur.

g. IAM rôle —Choisissez un IAM rôle existant ou créez-en un nouveau IAM pour accéder aux informations d'identification de votre référentiel et indexer le contenu.

> **Note**
>
> IAM les rôles utilisés pour les index ne peuvent pas être utilisés pour les sources de données. Si vous ne savez pas si un rôle existant est utilisé pour un index ou une FAQ, choisissez Créer un nouveau rôle pour éviter les erreurs.

h. Choisissez Suivant.

7. Sur la page Configurer les paramètres de synchronisation, entrez les informations suivantes :

a. Étendue de synchronisation : définissez des limites pour l'exploration de certains types de contenu, composants de page et chemins racines, et filtrez le contenu à l'aide de modèles d'expression regex.

i. Types de contenu : choisissez d'explorer uniquement les pages ou les ressources, ou les deux.

ii. (Facultatif) Configuration supplémentaire : configurez les paramètres suivants :

- Composants de page : noms spécifiques des composants de page. Le composant de page est un composant de page extensible conçu pour fonctionner avec l'éditeur de Adobe Experience Manager modèles et permet d'assembler les composants de page header/footer et de structure avec l'éditeur de modèles.
- Variations de fragments de contenu : noms spécifiques des variations de fragments de contenu. Les fragments de contenu vous permettent de concevoir, créer, organiser et publier du contenu indépendant de la page dans. Adobe Experience Manager Ils vous permettent de préparer du contenu prêt à être utilisé sur locations/over plusieurs canaux.
- Chemins racines : chemins racines vers un contenu spécifique.
- Modèles Regex : modèles d'expressions régulières permettant d'inclure ou d'exclure certaines pages et ressources.

b. Mode de synchronisation : choisissez la manière dont vous souhaitez mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation.

   - Synchronisation complète : Indexez tout le contenu fraîchement, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
   - Nouvelle synchronisation modifiée : Indexez uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
   - Synchronisation nouvelle, modifiée ou supprimée : Indexez uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

c. ID de fuseau horaire : si vous utilisez AEM On-Premise et que le fuseau horaire de votre serveur est différent de celui du connecteur ou de l'index Amazon Kendra AEM, vous pouvez spécifier le fuseau horaire du serveur afin de l'aligner sur le connecteur ou l'index AEM. Le fuseau horaire par défaut pour AEM On-Premise est le fuseau horaire du connecteur ou de l' Amazon Kendra index AEM. Le fuseau horaire par défaut pour AEM en tant que service cloud est l'heure moyenne de Greenwich.

d. Calendrier d'exécution de synchronisation, pour la fréquence : choisissez la fréquence à laquelle vous souhaitez synchroniser le contenu de votre source de données et mettre à jour votre index.

e. Choisissez Suivant.

8. Sur la page Définir les mappages de champs, entrez les informations suivantes :

   a. Sélectionnez parmi les champs de source de données par défaut Amazon Kendra générés que vous souhaitez mapper à votre index. Pour ajouter des champs de source de données personnalisés, créez un nom de champ d'index à mapper et le type de données du champ.

b. Choisissez Suivant.

9. Sur la page Réviser et créer, vérifiez que les informations saisies sont correctes, puis sélectionnez Ajouter une source de données. Vous pouvez également choisir de modifier vos informations à partir de cette page. Votre source de données apparaîtra sur la page Sources de données une fois que la source de données aura été ajoutée avec succès.

API

Pour vous connecter Amazon Kendra à Adobe Experience Manager

Vous devez spécifier un JSON du schéma de source de données à l'aide de l'TemplateConfigurationAPI. Vous devez fournir les informations suivantes :

- Source de données —Spécifiez le type de source de données tel que `AEM` lorsque vous utilisez le schéma TemplateConfigurationJSON. Spécifiez également la source de données `TEMPLATE` lorsque vous appelez l'CreateDataSourceAPI.
- URL de l'hôte AEM : spécifiez l'URL de l'Adobe Experience Managerhôte. Par exemple, si vous utilisez AEM On-Premise, vous devez inclure le nom d'hôte et le port :. https://hostname:port Ou, si vous utilisez AEM en tant que service cloud, vous pouvez utiliser l'URL de l'auteur : https://author-xxxxxx-xxxxxxx.adobeaemcloud.com.
- Mode de synchronisation : spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation. Vous pouvez choisir entre :
  - `FORCED_FULL_CRAWL`pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
  - `FULL_CRAWL`pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
  - `CHANGE_LOG`pour indexer uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

- Type d'authentification : spécifiez le type d'authentification que vous souhaitez utiliser, `Basic` soit`OAuth2`.
- Type AEM —Spécifiez le type Adobe Experience Manager que vous utilisez, soit`CLOUD`. `ON_PREMISE`
- Nom de ressource Amazon secret (ARN) : si vous souhaitez utiliser l'authentification de base pour AEM sur site ou dans le cloud, vous devez fournir un secret qui stocke vos informations d'authentification, à savoir votre nom d'utilisateur et votre mot de passe. Vous fournissez le Amazon Resource Name (ARN) d'un AWS Secrets Manager secret. Le secret est stocké dans une structure JSON avec les clés suivantes :

```
{
    "aemUrl": "Adobe Experience Manager On-Premise host URL",
    "username": "user name with admin permissions",
    "password": "password with admin permissions"
}
```

Si vous souhaitez utiliser l'authentification OAuth 2.0 pour AEM On-Premise, le secret est stocké dans une structure JSON avec les clés suivantes :

```
{
    "aemUrl": "Adobe Experience Manager host URL",
    "clientId": "client ID",
    "clientSecret": "client secret",
    "privateKey": "private key"
}
```

Si vous souhaitez utiliser l'authentification OAuth 2.0 pour AEM en tant que service cloud, le secret est stocké dans une structure JSON avec les clés suivantes :

```
{
    "clientId": "client ID",
    "clientSecret": "client secret",
    "privateKey": "private key",
    "orgId": "organization ID",
    "technicalAccountId": "technical account ID",
    "imsHost": "Adobe Identity Management System (IMS) host"
}
```

- IAM role —Spécifiez à quel `RoleArn` moment vous appelez `CreateDataSource` pour fournir à un IAM rôle les autorisations d'accéder à votre Secrets Manager secret et pour appeler le public requis APIs pour le connecteur Adobe Experience Manager et Amazon Kendra. Pour plus d'informations, consultez la section IAM Rôles des sources de données Adobe Experience Manager.

Vous pouvez également ajouter les fonctionnalités optionnelles suivantes :

- Virtual Private Cloud (VPC) **`VpcConfiguration`** —Spécifiez le moment de votre appel. `CreateDataSource` Pour de plus amples informations, veuillez consulter Configuration Amazon Kendra pour utiliser un Amazon VPC.
- ID de fuseau horaire : si vous utilisez AEM On-Premise et que le fuseau horaire de votre serveur est différent de celui du connecteur ou de l'index Amazon Kendra AEM, vous pouvez spécifier le fuseau horaire du serveur afin de l'aligner sur le connecteur ou l'index AEM.

  Le fuseau horaire par défaut pour AEM On-Premise est le fuseau horaire du connecteur ou de l' Amazon Kendra index AEM. Le fuseau horaire par défaut pour AEM en tant que service cloud est l'heure moyenne de Greenwich.

  Pour plus d'informations sur les fuseaux horaires pris en charge IDs, consultez le schéma Adobe Experience Manager JSON.
- Filtres d'inclusion et d'exclusion : spécifiez s'il faut inclure ou exclure certaines pages et ressources.

  > **Note**
  >
  > La plupart des sources de données utilisent des modèles d'expressions régulières, qui sont des modèles d'inclusion ou d'exclusion appelés filtres. Si vous spécifiez un filtre d'inclusion, seul le contenu correspondant au filtre d'inclusion est indexé. Tout document qui ne correspond pas au filtre d'inclusion n'est pas indexé. Si vous spécifiez un filtre d'inclusion et d'exclusion, les documents correspondant au filtre d'exclusion ne sont pas indexés, même s'ils correspondent au filtre d'inclusion.
- Identity Crawler : spécifiez s'il faut activer l'explorateur Amazon Kendra d'identité. Le robot d'exploration d'identité utilise les informations de la liste de contrôle d'accès (ACL) de vos documents pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Si vous disposez d'une ACL pour vos documents et que vous

choisissez de l'utiliser, vous pouvez également choisir d'activer le robot d'exploration Amazon Kendra d'identité pour configurer le filtrage des résultats de recherche par contexte utilisateur. Sinon, si le robot d'identification est désactivé, tous les documents peuvent être consultés publiquement. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'exploration d'identité est désactivé, vous pouvez également utiliser l'PutPrincipalMappingAPI pour télécharger les informations d'accès des utilisateurs et des groupes afin de filtrer le contexte utilisateur.

- Mappages de champs : choisissez de mapper les champs de votre source de données Adobe Experience Manager à vos champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.

> **Note**
>
> Le champ du corps du document ou l'équivalent du corps du document pour vos documents est requis pour Amazon Kendra effectuer une recherche dans vos documents. Vous devez associer le nom du champ du corps du document dans votre source de données au nom du champ d'index`_document_body`. Tous les autres champs sont facultatifs.

Pour obtenir la liste des autres clés JSON importantes à configurer, consultez le schéma du Adobe Experience Manager modèle.

Alfresco

> **Note**
>
> Le connecteur Alfresco reste entièrement compatible pour les clients existants jusqu'au 31 mai 2026. Bien que ce connecteur ne soit plus disponible pour les nouveaux utilisateurs, les utilisateurs actuels peuvent continuer à l'utiliser sans interruption. Nous faisons constamment évoluer notre portefeuille de connecteurs afin de proposer des solutions plus évolutives et personnalisables. Pour les intégrations futures, nous vous recommandons d'explorer le framework de connecteurs personnalisés Amazon Kendra [1], conçu pour prendre en charge un plus large éventail de cas d'utilisation en entreprise avec une flexibilité accrue.

Alfrescoest un service de gestion de contenu qui aide les clients à stocker et à gérer leur contenu. Vous pouvez l'utiliser Amazon Kendra pour indexer votre bibliothèque de Alfresco documents, votre wiki et votre blog.

Amazon Kendra prend en charge les services Alfresco sur site et Alfresco dans le cloud (plateforme en tant que service).

Vous pouvez vous connecter Amazon Kendra à votre source de Alfresco données à l'aide de la Amazon Kendra console ou de l'TemplateConfigurationAPI.

Pour résoudre les problèmes liés à votre connecteur de source de données Amazon Kendra Alfresco, consultez. Dépannage des sources de données

Rubriques

- Fonctionnalités prises en charge
- Prérequis
- Instructions de connexion
- En savoir plus

Fonctionnalités prises en charge

Amazon Kendra Alfrescole connecteur de source de données prend en charge les fonctionnalités suivantes :

- Mappages de champs
- Contrôle d'accès des utilisateurs
- Filtres d'inclusion/exclusion
- Synchronisation complète et incrémentielle du contenu
- OAuth Authentification 2.0 et de base
- Cloud privé virtuel (VPC)

Prérequis

Avant de pouvoir l'utiliser Amazon Kendra pour indexer votre source de données Alfresco, apportez ces modifications dans votre Alfresco fichier et. Comptes AWS

DansAlfresco, assurez-vous d'avoir :

- Vous avez copié l'URL de votre Alfresco référentiel et l'URL de votre application Web. Si vous souhaitez uniquement indexer un Alfresco site spécifique, copiez également l'ID du site.
- Notez vos informations Alfresco d'authentification, qui incluent un nom d'utilisateur et un mot de passe avec au moins des autorisations de lecture. Si vous souhaitez utiliser l'authentification OAuth 2.0, vous devez ajouter l'utilisateur au groupe Alfresco des administrateurs.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

- Facultatif : informations d'identification OAuth 2.0 configurées dansAlfresco. Les informations d'identification incluent l'identifiant du client, le secret du client et l'URL du jeton. Pour plus d'informations sur la configuration des clients sur Alfresco site, consultez la documentation Alfresco. Si vous utilisez Alfresco le Cloud (PaaS), vous devez contacter le support Hyland pour l'authentification Alfresco OAuth 2.0.
- Il est vérifié que chaque document est unique dans Alfresco et dans les autres sources de données que vous prévoyez d'utiliser pour le même index. Chaque source de données que vous souhaitez utiliser pour un index ne doit pas contenir le même document dans toutes les sources de données. IDs Les documents sont globaux par rapport à un index et doivent être uniques par index.

Dans votre Compte AWS, assurez-vous d'avoir :

- Création d'un Amazon Kendra index et, si vous utilisez l'API, notez l'ID de l'index.
- Vous avez créé un IAM rôle pour votre source de données et, si vous utilisez l'API, notez l'ARN du IAM rôle.

> **Note**
>
> Si vous modifiez votre type d'authentification et vos informations d'identification, vous devez mettre à jour votre IAM rôle pour accéder au bon identifiant AWS Secrets Manager secret.

- Stockez vos identifiants d'authentification Alfresco dans un AWS Secrets Manager secret et, si vous utilisez l'API, notez l'ARN du secret.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

Si vous n'avez pas de IAM rôle ou de secret existant, vous pouvez utiliser la console pour créer un nouveau IAM rôle et un nouveau Secrets Manager secret lorsque vous connectez votre source de données Alfresco à. Amazon Kendra Si vous utilisez l'API, vous devez fournir l'ARN d'un IAM rôle et d'un Secrets Manager secret existants, ainsi qu'un identifiant d'index.

Instructions de connexion

Pour vous connecter Amazon Kendra à votre source de données Alfresco, vous devez fournir les informations nécessaires à votre source de données Alfresco afin de Amazon Kendra pouvoir accéder à vos données. Si vous n'avez pas encore configuré Alfresco pour Amazon Kendra, consultez. Prérequis

Console

Pour vous connecter Amazon Kendra à Alfresco

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la.
2. Dans le volet de navigation de gauche, choisissez Index, puis choisissez l'index que vous souhaitez utiliser dans la liste des index.

   > **Note**
   >
   > Vous pouvez choisir de configurer ou de modifier vos paramètres de contrôle d'accès utilisateur dans les paramètres de l'index.

3. Sur la page de démarrage, choisissez Ajouter une source de données.

4. Sur la page Ajouter une source de données, choisissez le connecteur Alfresco, puis sélectionnez Ajouter un connecteur. Si vous utilisez la version 2 (le cas échéant), choisissez le connecteur Alfresco avec le tag « V2.0 ».
5. Sur la page Spécifier les détails de la source de données, entrez les informations suivantes :
   a. Dans Nom et description, pour Nom de la source de données : entrez le nom de votre source de données. Vous pouvez inclure des traits d'union, mais pas des espaces.
   b. (Facultatif) Description : entrez une description facultative pour votre source de données.
   c. Dans la langue par défaut : choisissez une langue pour filtrer vos documents pour l'index. Sauf indication contraire, la langue par défaut est l'anglais. La langue spécifiée dans les métadonnées du document remplace la langue sélectionnée.
   d. Dans Balises, pour Ajouter une nouvelle balise : incluez des balises facultatives pour rechercher et filtrer vos ressources ou suivre vos AWS coûts.
   e. Choisissez Suivant.
6. Sur la page Définir l'accès et la sécurité, entrez les informations suivantes :
   a. Alfrescotype —Choisissez si vous utilisez Alfresco sur place/serveur ou dans le Alfresco cloud (plate-forme en tant que service).
   b. URL du dépôt Alfresco : entrez l'URL de votre référentiel Alfresco. Par exemple, si vous utilisez Alfresco le Cloud (PaaS), l'URL du référentiel peut être https://company.alfrescocloud.com. Ou, si vous utilisez Alfresco On-Premises, l'URL du référentiel peut être. https://company-alfresco-instance.company-domain.suffix:port
   c. Application utilisateur Alfresco. URL —Entrez l'URL de votre interface Alfresco utilisateur. Vous pouvez obtenir l'URL du dépôt auprès de votre Alfresco administrateur. Par exemple, l'URL de l'interface utilisateur peut être https://example.com.
   d. Emplacement du certificat SSL : entrez le chemin d'accès au certificat SSL stocké dans un Amazon S3 compartiment. Vous l'utilisez pour vous connecter à Alfresco On-Premises avec une connexion SSL sécurisée.
   e. Autorisation : activez ou désactivez les informations de la liste de contrôle d'accès (ACL) pour vos documents, si vous disposez d'une ACL et que vous souhaitez l'utiliser pour le contrôle d'accès. L'ACL indique les documents auxquels les utilisateurs et les groupes peuvent accéder. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.

f. Authentification —Choisissez l'authentification de base ou l'authentification OAuth 2.0. Choisissez ensuite un Secrets Manager secret existant ou créez-en un nouveau pour stocker vos Alfresco informations d'identification. Si vous choisissez de créer un nouveau secret, une fenêtre AWS Secrets Manager secrète s'ouvre.

   Si vous avez choisi l'authentification de base, entrez le nom du secret, le nom Alfresco d'utilisateur et le mot de passe.

   Si vous avez choisi l'authentification OAuth 2.0, entrez un nom pour le secret, l'ID client, le secret client et l'URL du jeton.

g. Virtual Private Cloud (VPC) —Vous pouvez choisir d'utiliser un VPC. Dans ce cas, vous devez ajouter des sous-réseaux et des groupes de sécurité VPC.

h. Identity Crawler : spécifiez s'il faut activer l'explorateur Amazon Kendra d'identité. Le robot d'exploration d'identité utilise les informations de la liste de contrôle d'accès (ACL) de vos documents pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Si vous disposez d'une ACL pour vos documents et que vous choisissez de l'utiliser, vous pouvez également choisir d'activer le robot d'exploration Amazon Kendra d'identité pour configurer le filtrage des résultats de recherche par contexte utilisateur. Sinon, si le robot d'identification est désactivé, tous les documents peuvent être consultés publiquement. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'exploration d'identité est désactivé, vous pouvez également utiliser l'PutPrincipalMappingAPI pour télécharger les informations d'accès des utilisateurs et des groupes afin de filtrer le contexte utilisateur.

i. IAM rôle —Choisissez un IAM rôle existant ou créez-en un nouveau IAM pour accéder aux informations d'identification de votre référentiel et indexer le contenu.

> **Note**
>
> IAM les rôles utilisés pour les index ne peuvent pas être utilisés pour les sources de données. Si vous ne savez pas si un rôle existant est utilisé pour un index ou une FAQ, choisissez Créer un nouveau rôle pour éviter les erreurs.

j. Choisissez Suivant.

7. Sur la page Configurer les paramètres de synchronisation, entrez les informations suivantes :

   a. Étendue de synchronisation : définissez des limites pour l'exploration de certains contenus et filtrez le contenu à l'aide de modèles d'expression regex.

b. i. Contenu : choisissez d'explorer le contenu marqué d'un « Aspects »Alfresco, le contenu d'un Alfresco site spécifique ou le contenu de tous vos Alfresco sites.

ii. (Facultatif) Configuration supplémentaire : définissez les paramètres suivants :

- Inclure les commentaires : choisissez d'inclure les commentaires dans la bibliothèque de Alfresco documents et le blog.
- Modèles Regex : modèles d'expressions régulières permettant d'inclure ou d'exclure certains fichiers.

c. Mode de synchronisation : choisissez la manière dont vous souhaitez mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation.

- Synchronisation complète : Indexez tout le contenu fraîchement, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
- Synchronisation nouvelle, modifiée ou supprimée : Indexez uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

d. Dans Calendrier d'exécution de la synchronisation, pour Fréquence : choisissez la fréquence à laquelle vous souhaitez synchroniser le contenu de votre source de données et mettre à jour votre index.

e. Choisissez Suivant.

8. Sur la page Définir les mappages de champs, entrez les informations suivantes :

a. Sélectionnez parmi les champs de source de données par défaut Amazon Kendra générés que vous souhaitez mapper à votre index.

b. Pour ajouter des champs de source de données personnalisés, créez un nom de champ d'index à mapper et le type de données du champ.

c. Choisissez Suivant.

9. Sur la page Réviser et créer, vérifiez que les informations saisies sont correctes, puis sélectionnez Ajouter une source de données. Vous pouvez également choisir de modifier vos informations à partir de cette page. Votre source de données apparaîtra sur la page Sources de données une fois que la source de données aura été ajoutée avec succès.

API

Pour vous connecter Amazon Kendra à Alfresco

Vous devez spécifier un JSON du schéma de source de données à l'aide de l'TemplateConfigurationAPI. Vous devez fournir les informations suivantes :

- Source de données —Spécifiez le type de source de données tel que `ALFRESCO` lorsque vous utilisez le schéma TemplateConfigurationJSON. Spécifiez également la source de données `TEMPLATE` lorsque vous appelez l'CreateDataSourceAPI.
- AlfrescoID du site —Spécifiez l'ID du site Alfresco.
- AlfrescoURL du référentiel : spécifiez l'URL du Alfresco référentiel. Vous pouvez obtenir l'URL du dépôt auprès de votre Alfresco administrateur. Par exemple, si vous utilisez Alfresco le Cloud (PaaS), l'URL du référentiel peut être https://company.alfrescocloud.com. Ou, si vous utilisez Alfresco On-Premises, l'URL du référentiel peut être. https://company-alfresco-instance.company-domain.suffix:port
- AlfrescoURL de l'application Web —Spécifiez l'URL de Alfresco l'interface utilisateur. Vous pouvez obtenir l'URL du dépôt auprès de votre Alfresco administrateur. Par exemple, l'URL de l'interface utilisateur peut être https://example.com.
- Type d'authentification : spécifiez le type d'authentification que vous souhaitez utiliser, que ce soit `OAuth2` ou`Basic`.
- Alfrescotype —Spécifiez le type que Alfresco vous utilisez, que ce soit `PAAS` (Cloud/Plateforme en tant que service) ou `ON_PREM` (sur site).
- Nom de ressource Amazon secret (ARN) : si vous souhaitez utiliser l'authentification de base, vous devez fournir un secret qui stocke vos informations d'authentification, à savoir votre nom d'utilisateur et votre mot de passe. Vous fournissez le Amazon Resource Name (ARN) d'un AWS Secrets Manager secret. Le secret est stocké dans une structure JSON avec les clés suivantes :

```
{
    "username": "user name",
    "password": "password"
```

```
}
```

Si vous souhaitez utiliser l'authentification OAuth 2.0, le secret est stocké dans une structure JSON avec les clés suivantes :

```
{
    "clientId": "client ID",
    "clientSecret": "client secret",
    "tokenUrl": "token URL"
}
```

- IAM role —Spécifiez à quel `RoleArn` moment vous appelez `CreateDataSource` pour fournir à un IAM rôle les autorisations d'accéder à votre Secrets Manager secret et pour appeler le public requis APIs pour le connecteur Alfresco et. Amazon Kendra Pour plus d'informations, consultez la section IAM Rôles des sources de données Alfresco.

Vous pouvez également ajouter les fonctionnalités optionnelles suivantes :

- Virtual Private Cloud (VPC) **`VpcConfiguration`** —Spécifiez le moment de votre appel. `CreateDataSource` Pour de plus amples informations, veuillez consulter Configuration Amazon Kendra pour utiliser un Amazon VPC.
- Type de contenu : le type de contenu que vous souhaitez explorer, qu'il s'agisse du contenu marqué d'un « Aspects » dans le champ « Aspects »Alfresco, du contenu d'un Alfresco site spécifique ou du contenu de tous vos Alfresco sites. Vous pouvez également répertorier du contenu « Aspects » spécifique.
- Filtres d'inclusion et d'exclusion : spécifiez si vous souhaitez inclure ou exclure certains fichiers.

> **Note**
>
> La plupart des sources de données utilisent des modèles d'expressions régulières, qui sont des modèles d'inclusion ou d'exclusion appelés filtres. Si vous spécifiez un filtre d'inclusion, seul le contenu correspondant au filtre d'inclusion est indexé. Tout document qui ne correspond pas au filtre d'inclusion n'est pas indexé. Si vous spécifiez un filtre d'inclusion et d'exclusion, les documents correspondant au filtre d'exclusion ne sont pas indexés, même s'ils correspondent au filtre d'inclusion.

- Mode de synchronisation : spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de

données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation. Vous pouvez choisir entre :

- `FORCED_FULL_CRAWL`pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
- `FULL_CRAWL`pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

- Identity Crawler : spécifiez s'il faut activer l'explorateur Amazon Kendra d'identité. Le robot d'exploration d'identité utilise les informations de la liste de contrôle d'accès (ACL) de vos documents pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Si vous disposez d'une ACL pour vos documents et que vous choisissez de l'utiliser, vous pouvez également choisir d'activer le robot d'exploration Amazon Kendra d'identité pour configurer le filtrage des résultats de recherche par contexte utilisateur. Sinon, si le robot d'identification est désactivé, tous les documents peuvent être consultés publiquement. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'exploration d'identité est désactivé, vous pouvez également utiliser l'PutPrincipalMappingAPI pour télécharger les informations d'accès des utilisateurs et des groupes afin de filtrer le contexte utilisateur.
- Mappages de champs : choisissez de mapper les champs de votre source de données Alfresco à vos champs d'index. Amazon Kendra Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.

> **Note**
>
> Le champ du corps du document ou l'équivalent du corps du document pour vos documents est requis pour Amazon Kendra effectuer une recherche dans vos documents. Vous devez associer le nom du champ du corps du document dans votre source de données au nom du champ d'index`_document_body`. Tous les autres champs sont facultatifs.

Pour obtenir la liste des autres clés JSON importantes à configurer, consultez le schéma du Alfresco modèle.

En savoir plus

Pour en savoir plus sur l'intégration Amazon Kendra à votre source de données Alfresco, consultez :

- Recherchez du Alfresco contenu de manière intelligente à l'aide de Amazon Kendra

Aurora (MySQL)

> **Note**
>
> Aurora (MySQL)Le connecteur reste entièrement pris en charge pour les clients existants jusqu'au 31 mai 2026. Bien que ce connecteur ne soit plus disponible pour les nouveaux utilisateurs, les utilisateurs actuels peuvent continuer à l'utiliser sans interruption. Nous faisons constamment évoluer notre portefeuille de connecteurs afin de proposer des solutions plus évolutives et personnalisables. Pour les intégrations futures, nous vous recommandons d'explorer le framework de connecteurs personnalisés Amazon Kendra [1], conçu pour prendre en charge un plus large éventail de cas d'utilisation en entreprise avec une flexibilité accrue.

Aurora est un système de gestion de base de données relationnelle (RDBMS) conçu pour le cloud. Si vous êtes un Aurora utilisateur, vous pouvez l'utiliser Amazon Kendra pour indexer votre source de Aurora (MySQL) données. Le connecteur de source de Amazon Kendra Aurora (MySQL) données prend en charge Aurora MySQL 3 et MySQL 8.0 Aurora sans serveur.

Vous pouvez vous connecter Amazon Kendra à votre source de Aurora (MySQL) données à l'aide de la Amazon Kendra console et de l'TemplateConfigurationAPI.

Pour résoudre les problèmes liés à votre connecteur de source de Amazon Kendra Aurora (MySQL) données, consultezDépannage des sources de données.

Rubriques

- Fonctionnalités prises en charge
- Prérequis
- Instructions de connexion
- Remarques

Fonctionnalités prises en charge

- Mappages de champs
- Filtrage du contexte utilisateur
- Filtres d'inclusion/exclusion
- Synchronisation complète et incrémentielle du contenu
- Cloud privé virtuel (VPC)

Prérequis

Avant de pouvoir Amazon Kendra indexer votre source de Aurora (MySQL) données, apportez ces modifications à vos AWS comptes Aurora (MySQL) and.

DansAurora (MySQL), assurez-vous d'avoir :

- Notez le nom d'utilisateur et le mot de passe de votre base de données.

> **Important**
>
> Il est recommandé de fournir des informations d'identification Amazon Kendra de base de données en lecture seule.

- Vous avez copié l'URL, le port et l'instance de votre hôte de base de données. Vous pouvez trouver ces informations sur la Amazon RDS console.
- Il est vérifié que chaque document est unique dans Aurora (MySQL) et entre les autres sources de données que vous prévoyez d'utiliser pour le même index. Chaque source de données que vous souhaitez utiliser pour un index ne doit pas contenir le même document dans toutes les sources de données. IDs Les documents sont globaux par rapport à un index et doivent être uniques par index.

Dans votre Compte AWS, assurez-vous d'avoir :

- Création d'un Amazon Kendra index et, si vous utilisez l'API, notez l'ID de l'index.
- Vous avez créé un IAM rôle pour votre source de données et, si vous utilisez l'API, notez l'ARN du IAM rôle.

> **Note**
>
> Si vous modifiez votre type d'authentification et vos informations d'identification, vous devez mettre à jour votre IAM rôle pour accéder au bon identifiant AWS Secrets Manager secret.

- Stockez vos informations d'Aurora (MySQL)authentification dans un AWS Secrets Manager secret et, si vous utilisez l'API, notez l'ARN du secret.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

Si vous n'avez pas de IAM rôle ou de secret existant, vous pouvez utiliser la console pour créer un nouveau IAM rôle et un nouveau Secrets Manager secret lorsque vous connectez votre source de Aurora (MySQL) données à Amazon Kendra. Si vous utilisez l'API, vous devez fournir l'ARN d'un IAM rôle et d'un Secrets Manager secret existants, ainsi qu'un identifiant d'index.

Instructions de connexion

Pour vous connecter Amazon Kendra à votre source de Aurora (MySQL) données, vous devez fournir les détails de vos Aurora (MySQL) informations d'identification afin de Amazon Kendra pouvoir accéder à vos données. Si vous n'avez pas encore configuré Aurora (MySQL) pour Amazon Kendra voirPrérequis.

Console

Pour vous connecter Amazon Kendra à Aurora (MySQL)

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la.
2. Dans le volet de navigation de gauche, choisissez Index, puis choisissez l'index que vous souhaitez utiliser dans la liste des index.

> **Note**
>
> Vous pouvez choisir de configurer ou de modifier vos paramètres de contrôle d'accès utilisateur dans les paramètres de l'index.

3. Sur la page de démarrage, choisissez Ajouter une source de données.
4. Sur la page Ajouter une source de données, choisissez Aurora (MySQL)connecteur, puis sélectionnez Ajouter un connecteur. Si vous utilisez la version 2 (le cas échéant), choisissez le Aurora (MySQL)connecteur avec la balise « V2.0 ».
5. Sur la page Spécifier les détails de la source de données, entrez les informations suivantes :
   a. Dans Nom et description, pour Nom de la source de données : entrez le nom de votre source de données. Vous pouvez inclure des traits d'union, mais pas des espaces.
   b. (Facultatif) Description : entrez une description facultative pour votre source de données.
   c. Dans la langue par défaut : choisissez une langue pour filtrer vos documents pour l'index. Sauf indication contraire, la langue par défaut est l'anglais. La langue spécifiée dans les métadonnées du document remplace la langue sélectionnée.
   d. Dans Balises, pour Ajouter une nouvelle balise : incluez des balises facultatives pour rechercher et filtrer vos ressources ou suivre vos AWS coûts.
   e. Choisissez Suivant.
6. Sur la page Définir l'accès et la sécurité, entrez les informations suivantes :
   a. Dans Source, entrez les informations suivantes :
   b. Hôte : entrez l'URL de l'hôte de la base de données, par exemple :`http://`*`instance URL`*`.`*`region`*`.rds.amazonaws.com`.
   c. Port — Entrez le port de base de données, par exemple,`5432`.
   d. Instance — Entrez l'instance de base de données.
   e. Dans Authentification, entrez les informations suivantes :
      - AWS Secrets Manager secret —Choisissez un secret existant ou créez-en un nouveau Secrets Manager pour stocker vos informations Aurora (MySQL) d'authentification. Si vous choisissez de créer un nouveau secret, une fenêtre AWS Secrets Manager secrète s'ouvre.

A. Entrez les informations suivantes dans la fenêtre Créer un AWS Secrets Manager secret :

I. Nom secret : le nom de votre secret. Le préfixe « AmazonKendra - Aurora (MySQL) - » est automatiquement ajouté à votre nom secret.

II. Pour le nom d'utilisateur et le mot de passe de la base de données, entrez les valeurs d'identification d'authentification que vous avez copiées depuis votre base de données.

B. Choisissez Enregistrer.

f. Virtual Private Cloud (VPC) —Vous pouvez choisir d'utiliser un VPC. Dans ce cas, vous devez ajouter des sous-réseaux et des groupes de sécurité VPC.

g. IAM rôle —Choisissez un IAM rôle existant ou créez-en un nouveau IAM pour accéder aux informations d'identification de votre référentiel et indexer le contenu.

> **Note**
>
> IAM les rôles utilisés pour les index ne peuvent pas être utilisés pour les sources de données. Si vous ne savez pas si un rôle existant est utilisé pour un index ou une FAQ, choisissez Créer un nouveau rôle pour éviter les erreurs.

h. Choisissez Suivant.

7. Sur la page Configurer les paramètres de synchronisation, entrez les informations suivantes :

a. Dans le champ d'application de la synchronisation, choisissez l'une des options suivantes :

- Requête SQL —Entrez des instructions de requête SQL telles que les opérations SELECT et JOIN. Les requêtes SQL doivent être inférieures à 32 Ko Les requêtes SQL doivent être inférieures à 32 Ko et ne pas contenir de points-virgules (;). Amazon Kendra analysera tout le contenu de la base de données correspondant à votre requête.
- Colonne clé primaire : indiquez la clé primaire pour la table de base de données. Cela permet d'identifier une table au sein de votre base de données.
- Colonne de titre : indiquez le nom de la colonne de titre du document dans votre table de base de données.

- Colonne du corps : indiquez le nom de la colonne du corps du document dans votre table de base de données.

b. Dans Configuration supplémentaire — facultatif, choisissez l'une des options suivantes pour synchroniser un contenu spécifique au lieu de synchroniser tous les fichiers :

- Colonnes détectant les modifications : entrez le nom des colonnes qui Amazon Kendra seront utilisées pour détecter les modifications de contenu. Amazon Kendra réindexera le contenu en cas de modification de l'une de ces colonnes.
- IDsColonne des utilisateurs : entrez le nom de la colonne contenant l'utilisateur devant être autorisé IDs à accéder au contenu.
- Colonne Groupes : entrez le nom de la colonne contenant les groupes autorisés à accéder au contenu.
- URLsColonne source —Entrez le nom de la colonne contenant la source URLs à indexer.
- Colonne horodatage : entrez le nom de la colonne contenant les horodatages. Amazon Kendra utilise les informations d'horodatage pour détecter les modifications apportées à votre contenu et synchroniser uniquement le contenu modifié.
- Colonne fuseaux horaires : entrez le nom de la colonne contenant les fuseaux horaires du contenu à analyser.
- Format d'horodatage : entrez le nom de la colonne contenant les formats d'horodatage à utiliser pour détecter les modifications de contenu et resynchroniser votre contenu.

c. Mode de synchronisation : choisissez la manière dont vous souhaitez mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation.

- Synchronisation complète : Indexez tout le contenu fraîchement, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
- Nouvelle synchronisation modifiée : Indexez uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour

suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

- Synchronisation nouvelle, modifiée ou supprimée : Indexez uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

d. Dans Synchroniser le calendrier d'exécution, pour Fréquence : à quelle fréquence Amazon Kendra sera synchronisée avec votre source de données.

e. Choisissez Suivant.

8. Sur la page Définir les mappages de champs, entrez les informations suivantes :

a. Sélectionnez parmi les champs de source de données par défaut générés (Document IDs, Titres des documents et Source URLs) que vous souhaitez Amazon Kendra mapper à l'index.

b. Ajouter un champ —Pour ajouter des champs de source de données personnalisés afin de créer un nom de champ d'index à mapper et le type de données du champ.

c. Choisissez Suivant.

9. Sur la page Réviser et créer, vérifiez que les informations saisies sont correctes, puis sélectionnez Ajouter une source de données. Vous pouvez également choisir de modifier vos informations à partir de cette page. Votre source de données apparaîtra sur la page Sources de données une fois qu'elle aura été ajoutée avec succès.

API

Pour vous connecter Amazon Kendra à Aurora (MySQL)

Vous devez spécifier les éléments suivants à l'aide de l'TemplateConfigurationAPI :

- Source de données —Spécifiez le type de source de données tel que `JDBC` lorsque vous utilisez le schéma TemplateConfigurationJSON. Spécifiez également la source de données `TEMPLATE` lorsque vous appelez l'CreateDataSourceAPI.
- Type de base de données —Vous devez spécifier le type de base de données sous `mySql` la forme.

- Requête SQL : spécifiez des instructions de requête SQL telles que les opérations SELECT et JOIN. La taille des requêtes SQL doit être inférieure à 32 Ko. Amazon Kendra analysera tout le contenu de la base de données correspondant à votre requête.
- Mode de synchronisation : spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation. Vous pouvez choisir entre :
  - `FORCED_FULL_CRAWL`pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
  - `FULL_CRAWL`pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
  - `CHANGE_LOG`pour indexer uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
- Nom de ressource Amazon (ARN) secret : indiquez le nom de ressource Amazon (ARN) d'un Secrets Manager secret contenant les informations d'authentification que vous avez créées dans votre Aurora (MySQL) compte. Le secret est stocké dans une structure JSON avec les clés suivantes :

```
{
    "user name": "database user name",
    "password": "password"
}
```

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les

informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

- IAM role —Spécifiez à quel `RoleArn` moment vous appelez `CreateDataSource` pour fournir à un IAM rôle les autorisations d'accéder à votre Secrets Manager secret et d'appeler le public requis APIs pour le Aurora (MySQL) connecteur et Amazon Kendra. Pour plus d'informations, consultez la section IAM Rôles des sources de Aurora (MySQL) données.

Vous pouvez également ajouter les fonctionnalités optionnelles suivantes :

- Virtual Private Cloud (VPC) **`VpcConfiguration`** —Spécifiez le moment de votre appel. `CreateDataSource` Pour de plus amples informations, veuillez consulter Configuration Amazon Kendra pour utiliser un Amazon VPC.
- Filtres d'inclusion et d'exclusion : vous pouvez spécifier si vous souhaitez inclure un contenu spécifique en utilisant l'utilisateur IDs, les groupes, la source URLs, les horodatages et les fuseaux horaires.
- Filtrage du contexte utilisateur et contrôle d'accèsAmazon Kendra : analyse la liste de contrôle d'accès (ACL) de vos documents, si vous disposez d'une ACL pour vos documents. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.
- Mappages de champs : choisissez de mapper les champs de votre source de Aurora (MySQL) données à vos champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.

> **Note**
>
> Le champ du corps du document ou l'équivalent du corps du document pour vos documents est requis pour Amazon Kendra effectuer une recherche dans vos documents. Vous devez associer le nom du champ du corps du document dans votre source de données au nom du champ d'index`_document_body`. Tous les autres champs sont facultatifs.

Pour obtenir la liste des autres clés JSON importantes à configurer, consultezAurora Schéma de modèle (MySQL).

Remarques

- Les lignes de base de données supprimées ne seront pas suivies lors de la Amazon Kendra vérification du contenu mis à jour.
- La taille des noms et des valeurs des champs d'une ligne de votre base de données ne peut pas dépasser 400 Ko.
- Si la source de données de votre base de données contient une grande quantité de données et que vous ne souhaitez pas Amazon Kendra indexer tout le contenu de votre base de données après la première synchronisation, vous pouvez choisir de ne synchroniser que les documents nouveaux, modifiés ou supprimés.
- Il est recommandé de fournir des informations d'identification Amazon Kendra de base de données en lecture seule.
- Il est recommandé d'éviter d'ajouter des tableaux contenant des données sensibles ou des informations personnelles identifiables (PII).

Aurora (PostgreSQL)

> **Note**
>
> Aurora (PostgreSQL)Le connecteur reste entièrement pris en charge pour les clients existants jusqu'au 31 mai 2026. Bien que ce connecteur ne soit plus disponible pour les nouveaux utilisateurs, les utilisateurs actuels peuvent continuer à l'utiliser sans interruption. Nous faisons constamment évoluer notre portefeuille de connecteurs afin de proposer des solutions plus évolutives et personnalisables. Pour les intégrations futures, nous vous recommandons d'explorer le framework de connecteurs personnalisés Amazon Kendra [1], conçu pour prendre en charge un plus large éventail de cas d'utilisation en entreprise avec une flexibilité accrue.

Aurora est un système de gestion de base de données relationnelle (RDBMS) conçu pour le cloud. Si vous êtes un Aurora utilisateur, vous pouvez l'utiliser Amazon Kendra pour indexer votre source de Aurora (PostgreSQL) données. Le connecteur de source de Amazon Kendra Aurora (PostgreSQL) données prend en charge Aurora PostgreSQL 1.

Vous pouvez vous connecter Amazon Kendra à votre source de Aurora (PostgreSQL) données à l'aide de la Amazon Kendra console et de l'TemplateConfigurationAPI.

Pour résoudre les problèmes liés à votre connecteur de source de Amazon Kendra Aurora (PostgreSQL) données, consultezDépannage des sources de données.

Rubriques

- Fonctionnalités prises en charge
- Prérequis
- Instructions de connexion
- Remarques

Fonctionnalités prises en charge

- Mappages de champs
- Filtrage du contexte utilisateur
- Filtres d'inclusion/exclusion
- Synchronisation complète et incrémentielle du contenu
- Cloud privé virtuel (VPC)

Prérequis

Avant de pouvoir Amazon Kendra indexer votre source de Aurora (PostgreSQL) données, apportez ces modifications à vos AWS comptes Aurora (PostgreSQL) and.

DansAurora (PostgreSQL), assurez-vous d'avoir :

- Notez le nom d'utilisateur et le mot de passe de votre base de données.

> ⚠ Important
>
> Il est recommandé de fournir des informations d'identification Amazon Kendra de base de données en lecture seule.

- Vous avez copié l'URL, le port et l'instance de votre hôte de base de données.
- Il est vérifié que chaque document est unique dans Aurora (PostgreSQL) et entre les autres sources de données que vous prévoyez d'utiliser pour le même index. Chaque source de données que vous souhaitez utiliser pour un index ne doit pas contenir le même document dans toutes

les sources de données. IDs Les documents sont globaux par rapport à un index et doivent être uniques par index.

Dans votre Compte AWS, assurez-vous d'avoir :

- Création d'un Amazon Kendra index et, si vous utilisez l'API, notez l'ID de l'index.
- Vous avez créé un IAM rôle pour votre source de données et, si vous utilisez l'API, notez l'ARN du IAM rôle.

> **Note**
>
> Si vous modifiez votre type d'authentification et vos informations d'identification, vous devez mettre à jour votre IAM rôle pour accéder au bon identifiant AWS Secrets Manager secret.

- Stockez vos informations d'Aurora (PostgreSQL)authentification dans un AWS Secrets Manager secret et, si vous utilisez l'API, notez l'ARN du secret.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

Si vous n'avez pas de IAM rôle ou de secret existant, vous pouvez utiliser la console pour créer un nouveau IAM rôle et un nouveau Secrets Manager secret lorsque vous connectez votre source de Aurora (PostgreSQL) données à Amazon Kendra. Si vous utilisez l'API, vous devez fournir l'ARN d'un IAM rôle et d'un Secrets Manager secret existants, ainsi qu'un identifiant d'index.

Instructions de connexion

Pour vous connecter Amazon Kendra à votre source de Aurora (PostgreSQL) données, vous devez fournir les détails de vos Aurora (PostgreSQL) informations d'identification afin de Amazon Kendra pouvoir accéder à vos données. Si vous n'avez pas encore configuré Aurora (PostgreSQL) pour Amazon Kendra voirPrérequis.

Console

Pour vous connecter Amazon Kendra à Aurora (PostgreSQL)

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la.
2. Dans le volet de navigation de gauche, choisissez Index, puis choisissez l'index que vous souhaitez utiliser dans la liste des index.

> **Note**
>
> Vous pouvez choisir de configurer ou de modifier vos paramètres de contrôle d'accès utilisateur dans les paramètres de l'index.

3. Sur la page de démarrage, choisissez Ajouter une source de données.
4. Sur la page Ajouter une source de données, choisissez Aurora (PostgreSQL)connecteur, puis sélectionnez Ajouter un connecteur. Si vous utilisez la version 2 (le cas échéant), choisissez le Aurora (PostgreSQL)connecteur avec la balise « V2.0 ».
5. Sur la page Spécifier les détails de la source de données, entrez les informations suivantes :
   a. Dans Nom et description, pour Nom de la source de données : entrez le nom de votre source de données. Vous pouvez inclure des tirets, mais pas des espaces.
   b. (Facultatif) Description : entrez une description facultative pour votre source de données.
   c. Dans la langue par défaut : choisissez une langue pour filtrer vos documents pour l'index. Sauf indication contraire, la langue par défaut est l'anglais. La langue spécifiée dans les métadonnées du document remplace la langue sélectionnée.
   d. Dans Balises, pour Ajouter une nouvelle balise : incluez des balises facultatives pour rechercher et filtrer vos ressources ou suivre vos AWS coûts.
   e. Choisissez Suivant.
6. Sur la page Définir l'accès et la sécurité, entrez les informations suivantes :
   a. Dans Source, entrez les informations suivantes :
   b. Hôte : entrez l'URL de l'hôte de la base de données, par exemple :`http://`*`instance URL`*`.`*`region`*`.rds.amazonaws.com`.
   c. Port — Entrez le port de base de données, par exemple,`5432`.
   d. Instance — Entrez l'instance de base de données, par exemple`postgres`.

e. Activer l'emplacement du certificat SSL : choisissez d'entrer le Amazon S3 chemin d'accès à votre fichier de certificat SSL.

f. Dans Authentification, entrez les informations suivantes :

- AWS Secrets Manager secret —Choisissez un secret existant ou créez-en un nouveau Secrets Manager pour stocker vos informations Aurora (PostgreSQL) d'authentification. Si vous choisissez de créer un nouveau secret, une fenêtre AWS Secrets Manager secrète s'ouvre.

  A. Entrez les informations suivantes dans la fenêtre Créer un AWS Secrets Manager secret :

     I. Nom secret : le nom de votre secret. Le préfixe « AmazonKendra - Aurora (PostgreSQL) - » est automatiquement ajouté à votre nom secret.

     II. Pour le nom d'utilisateur et le mot de passe de la base de données, entrez les valeurs d'identification d'authentification que vous avez copiées depuis votre base de données.

  B. Choisissez Enregistrer.

g. Virtual Private Cloud (VPC) —Vous pouvez choisir d'utiliser un VPC. Dans ce cas, vous devez ajouter des sous-réseaux et des groupes de sécurité VPC.

h. IAM rôle —Choisissez un IAM rôle existant ou créez-en un nouveau IAM pour accéder aux informations d'identification de votre référentiel et indexer le contenu.

> **Note**
>
> IAM les rôles utilisés pour les index ne peuvent pas être utilisés pour les sources de données. Si vous ne savez pas si un rôle existant est utilisé pour un index ou une FAQ, choisissez Créer un nouveau rôle pour éviter les erreurs.

i. Choisissez Suivant.

7. Sur la page Configurer les paramètres de synchronisation, entrez les informations suivantes :

a. Dans le champ d'application de la synchronisation, choisissez l'une des options suivantes :

- Requête SQL —Entrez des instructions de requête SQL telles que les opérations SELECT et JOIN. Les requêtes SQL doivent être inférieures à 32 Ko Les requêtes

SQL doivent être inférieures à 32 Ko et ne pas contenir de points-virgules (;). Amazon Kendra analysera tout le contenu de la base de données correspondant à votre requête.

- Colonne clé primaire : indiquez la clé primaire pour la table de base de données. Cela identifie une table au sein de votre base de données.
- Colonne de titre : indiquez le nom de la colonne de titre du document dans votre table de base de données.
- Colonne du corps : indiquez le nom de la colonne du corps du document dans votre table de base de données.

b. Dans Configuration supplémentaire — facultatif, choisissez l'une des options suivantes pour synchroniser un contenu spécifique au lieu de synchroniser tous les fichiers :

- Colonnes détectant les modifications : entrez le nom des colonnes qui Amazon Kendra seront utilisées pour détecter les modifications de contenu. Amazon Kendra réindexera le contenu en cas de modification de l'une de ces colonnes.
- IDsColonne des utilisateurs : entrez le nom de la colonne contenant l'utilisateur devant être autorisé IDs à accéder au contenu.
- Colonne Groupes : entrez le nom de la colonne contenant les groupes autorisés à accéder au contenu.
- URLsColonne source —Entrez le nom de la colonne contenant la source URLs à indexer.
- Colonne horodatage : entrez le nom de la colonne contenant les horodatages. Amazon Kendra utilise les informations d'horodatage pour détecter les modifications apportées à votre contenu et synchroniser uniquement le contenu modifié.
- Colonne fuseaux horaires : entrez le nom de la colonne contenant les fuseaux horaires du contenu à analyser.
- Format d'horodatage : entrez le nom de la colonne contenant les formats d'horodatage à utiliser pour détecter les modifications de contenu et resynchroniser votre contenu.

c. Mode de synchronisation : choisissez la manière dont vous souhaitez mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation.

- Synchronisation complète : Indexez tout le contenu fraîchement, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
- Nouvelle synchronisation modifiée : Indexez uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
- Synchronisation nouvelle, modifiée ou supprimée : Indexez uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

d. Dans Synchroniser le calendrier d'exécution, pour Fréquence : à quelle fréquence Amazon Kendra sera synchronisée avec votre source de données.

e. Choisissez Suivant.

8. Sur la page Définir les mappages de champs, entrez les informations suivantes :

   a. Sélectionnez parmi les champs de source de données par défaut générés (Document IDs, Titres des documents et Source URLs) que vous souhaitez Amazon Kendra mapper à l'index.

   b. Ajouter un champ —Pour ajouter des champs de source de données personnalisés afin de créer un nom de champ d'index à mapper et le type de données du champ.

   c. Choisissez Suivant.

9. Sur la page Réviser et créer, vérifiez que les informations saisies sont correctes, puis sélectionnez Ajouter une source de données. Vous pouvez également choisir de modifier vos informations depuis cette page. Votre source de données apparaîtra sur la page Sources de données une fois que la source de données aura été ajoutée avec succès.

API

Pour vous connecter Amazon Kendra à Aurora (PostgreSQL)

Vous devez spécifier les éléments suivants à l'aide de l'TemplateConfigurationAPI :

- Source de données —Spécifiez le type de source de données tel que `JDBC` lorsque vous utilisez le schéma TemplateConfigurationJSON. Spécifiez également la source de données `TEMPLATE` lorsque vous appelez l'CreateDataSourceAPI.
- Type de base de données —Vous devez spécifier le type de base de données sous `postgresql` la forme.
- Requête SQL : spécifiez des instructions de requête SQL telles que les opérations SELECT et JOIN. La taille des requêtes SQL doit être inférieure à 32 Ko. Amazon Kendra analysera tout le contenu de la base de données correspondant à votre requête.
- Mode de synchronisation : spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation. Vous pouvez choisir entre :
  - `FORCED_FULL_CRAWL`pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
  - `FULL_CRAWL`pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
  - `CHANGE_LOG`pour indexer uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
- Nom de ressource Amazon (ARN) secret : indiquez le nom de ressource Amazon (ARN) d'un Secrets Manager secret contenant les informations d'authentification que vous avez créées dans votre Aurora (PostgreSQL) compte. Le secret est stocké dans une structure JSON avec les clés suivantes :

```
{
    "user name": "database user name",
    "password": "password"
}
```

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

- IAM role —Spécifiez à quel `RoleArn` moment vous appelez `CreateDataSource` pour fournir à un IAM rôle les autorisations d'accéder à votre Secrets Manager secret et d'appeler le public requis APIs pour le Aurora (PostgreSQL) connecteur et Amazon Kendra. Pour plus d'informations, consultez la section IAM Rôles des sources de Aurora (PostgreSQL) données.

Vous pouvez également ajouter les fonctionnalités optionnelles suivantes :

- Virtual Private Cloud (VPC) **`VpcConfiguration`** —Spécifiez le moment de votre appel. `CreateDataSource` Pour de plus amples informations, veuillez consulter Configuration Amazon Kendra pour utiliser un Amazon VPC.
- Filtres d'inclusion et d'exclusion : vous pouvez spécifier si vous souhaitez inclure un contenu spécifique en utilisant l'utilisateur IDs, les groupes, la source URLs, les horodatages et les fuseaux horaires.
- Filtrage du contexte utilisateur et contrôle d'accèsAmazon Kendra : analyse la liste de contrôle d'accès (ACL) de vos documents, si vous disposez d'une ACL pour vos documents. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.
- Mappages de champs : choisissez de mapper les champs de votre source de Aurora (PostgreSQL) données à vos champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.

> **Note**
>
> Le champ du corps du document ou l'équivalent du corps du document pour vos documents est requis pour Amazon Kendra effectuer une recherche dans vos documents. Vous devez associer le nom du champ du corps du document dans votre

source de données au nom du champ d'index_document_body. Tous les autres champs sont facultatifs.

Pour obtenir la liste des autres clés JSON importantes à configurer, consultezAurora Schéma de modèle (PostgreSQL).

Remarques

- Les lignes de base de données supprimées ne seront pas suivies lors de la Amazon Kendra vérification du contenu mis à jour.
- La taille des noms et des valeurs des champs d'une ligne de votre base de données ne peut pas dépasser 400 Ko.
- Si la source de données de votre base de données contient une grande quantité de données et que vous ne souhaitez pas Amazon Kendra indexer tout le contenu de votre base de données après la première synchronisation, vous pouvez choisir de ne synchroniser que les documents nouveaux, modifiés ou supprimés.
- Il est recommandé de fournir des informations d'identification Amazon Kendra de base de données en lecture seule.
- Il est recommandé d'éviter d'ajouter des tableaux contenant des données sensibles ou des informations personnelles identifiables (PII).

Amazon FSx (Fenêtres)

Amazon FSx (Windows) est un système de serveur de fichiers entièrement géré basé sur le cloud qui offre des fonctionnalités de stockage partagé. Si vous êtes un utilisateur Amazon FSx (Windows), vous pouvez l'utiliser Amazon Kendra pour indexer votre source de données Amazon FSx (Windows).

> **Note**
>
> Amazon Kendra prend désormais en charge un connecteur mis à niveau Amazon FSx (Windows).
> La console a été automatiquement mise à niveau pour vous. Tous les nouveaux connecteurs que vous créez sur la console utiliseront l'architecture mise à niveau. Si

vous utilisez l'API, vous devez désormais utiliser l'TemplateConfigurationobjet au lieu de l'`FSxConfiguration`objet pour configurer votre connecteur.
Les connecteurs configurés à l'aide de l'ancienne console et de l'ancienne architecture d'API continueront de fonctionner tels qu'ils ont été configurés. Toutefois, vous ne pourrez ni les modifier ni les mettre à jour. Si vous souhaitez modifier ou mettre à jour la configuration de votre connecteur, vous devez créer un nouveau connecteur.
Nous vous recommandons de migrer le flux de travail de votre connecteur vers la version mise à niveau. Support pour les connecteurs configurés à l'aide de l'ancienne architecture devrait prendre fin en juin 2024.

Vous pouvez vous connecter Amazon Kendra à votre source de données Amazon FSx (Windows) à l'aide de la Amazon Kendra console ou du TemplateConfigurationAPI.

Pour résoudre les problèmes liés à votre connecteur de source de données Amazon Kendra Amazon FSx (Windows), consultezDépannage des sources de données.

Rubriques

- Fonctionnalités prises en charge
- Prérequis
- Instructions de connexion
- En savoir plus

Fonctionnalités prises en charge

Amazon Kendra Amazon FSx Le connecteur de source de données (Windows) prend en charge les fonctionnalités suivantes :

- Mappages de champs
- Contrôle d'accès utilisateur
- Exploration de l'identité des utilisateurs
- Filtres d'inclusion et d'exclusion
- Synchronisation complète et incrémentielle du contenu
- Cloud privé virtuel (VPC)

Prérequis

Avant de pouvoir Amazon Kendra indexer votre source de données Amazon FSx (Windows), vérifiez les détails de votre Amazon FSx (Windows) et Comptes AWS.

Pour Amazon FSx (Windows), assurez-vous d'avoir :

- Configurez Amazon FSx (Windows) avec des autorisations de lecture et de montage.
- Vous avez noté l'identifiant de votre système de fichiers. Vous trouverez l'ID de votre système de fichiers sur le tableau de bord des systèmes de fichiers de la console Amazon FSx (Windows).
- Configuré un cloud privé virtuel en utilisant l' Amazon VPC emplacement de votre système de fichiers Amazon FSx (Windows).
- Vous avez noté vos informations d'authentification Amazon FSx (Windows) pour un Active Directory compte utilisateur. Cela inclut votre nom d'utilisateur Active Directory avec votre nom de domaine DNS (par exemple, user@corp.example.com) et votre mot de passe.

  > **Note**
  >
  > Utilisez uniquement les informations d'identification nécessaires au fonctionnement du connecteur. N'utilisez pas d'informations d'identification privilégiées telles que celles d'administrateur de domaine.

  > **Note**
  >
  > Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

- Il est vérifié que chaque document est unique dans Amazon FSx (Windows) et dans les autres sources de données que vous prévoyez d'utiliser pour le même index. Chaque source de données que vous souhaitez utiliser pour un index ne doit pas contenir le même document dans toutes les sources de données. IDs Les documents sont globaux par rapport à un index et doivent être uniques par index.

Dans votre Compte AWS, assurez-vous d'avoir :

- Création d'un Amazon Kendra index et, si vous utilisez l'API, notez l'ID de l'index.
- Vous avez créé un IAM rôle pour votre source de données et, si vous utilisez l'API, notez l'ARN du IAM rôle.

> **Note**
>
> Si vous modifiez votre type d'authentification et vos informations d'identification, vous devez mettre à jour votre IAM rôle pour accéder au bon identifiant AWS Secrets Manager secret.

- Stockez vos informations d'authentification Amazon FSx (Windows) dans un AWS Secrets Manager secret et, si vous utilisez l'API, notez l'ARN du secret.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

Si vous n'avez pas de IAM rôle ou de secret existant, vous pouvez utiliser la console pour créer un nouveau IAM rôle et un nouveau Secrets Manager secret lorsque vous connectez votre source de données Amazon FSx (Windows) à Amazon Kendra. Si vous utilisez l'API, vous devez fournir l'ARN d'un IAM rôle et d'un Secrets Manager secret existants, ainsi qu'un identifiant d'index.

Instructions de connexion

Pour vous connecter Amazon Kendra à votre source de données Amazon FSx (Windows), vous devez fournir les informations nécessaires sur votre source de données Amazon FSx (Windows) afin de Amazon Kendra pouvoir accéder à vos données. Si vous n'avez pas encore configuré Amazon FSx (Windows) pour Amazon Kendra, consultezPrérequis.

Console

Pour vous connecter Amazon Kendra à votre système de fichiers Amazon FSx (Windows)

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la.
2. Dans le volet de navigation de gauche, choisissez Index, puis choisissez l'index que vous souhaitez utiliser dans la liste des index.

> **Note**
>
> Vous pouvez choisir de configurer ou de modifier vos paramètres de contrôle d'accès utilisateur dans les paramètres de l'index.

3. Sur la page de démarrage, choisissez Ajouter une source de données.
4. Sur la page Ajouter une source de données, choisissez le connecteur Amazon FSx (Windows), puis choisissez Ajouter un connecteur. Si vous utilisez la version 2 (le cas échéant), choisissez le connecteur Amazon FSx (Windows) avec la balise « V2.0 ».
5. Sur la page Spécifier les détails de la source de données, entrez les informations suivantes :

   a. Dans Nom et description, pour Nom de la source de données : entrez le nom de votre source de données. Vous pouvez inclure des tirets, mais pas des espaces.

   b. (Facultatif) Description : entrez une description facultative pour votre source de données.

   c. Dans la langue par défaut : choisissez une langue pour filtrer vos documents pour l'index. Sauf indication contraire, la langue par défaut est l'anglais. La langue spécifiée dans les métadonnées du document remplace la langue sélectionnée.

   d. Dans Balises, pour Ajouter une nouvelle balise : incluez des balises facultatives pour rechercher et filtrer vos ressources ou suivre vos AWS coûts.

   e. Choisissez Suivant.

6. Sur la page Définir l'accès et la sécurité, entrez les informations suivantes :

   a. Amazon FSx ID du système de fichiers (Windows) —Sélectionnez dans le menu déroulant l'ID de votre système de fichiers existant, extrait de Amazon FSx (Windows). Vous pouvez également créer un système de fichiers Amazon FSx (Windows). Vous trouverez l'ID de votre système de fichiers sur le tableau de bord des systèmes de fichiers de la console Amazon FSx (Windows).

   b. Autorisation : activez ou désactivez les informations de la liste de contrôle d'accès (ACL) pour vos documents, si vous disposez d'une ACL et que vous souhaitez l'utiliser pour le contrôle d'accès. L'ACL indique les documents auxquels les utilisateurs et les groupes peuvent accéder. Les informations ACL sont utilisées pour filtrer les résultats de

recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.

c. Authentification : choisissez un AWS Secrets Manager secret existant ou créez-en un nouveau pour stocker les informations d'identification de votre système de fichiers. Si vous choisissez de créer un nouveau secret, une fenêtre AWS Secrets Manager secrète s'ouvre.

   Fournissez un secret qui stocke vos informations d'authentification, à savoir votre nom d'utilisateur et votre mot de passe. Le nom d'utilisateur doit inclure votre nom de domaine DNS. Par exemple, user@corp.example.com.

   Enregistrez et ajoutez votre secret.

d. Virtual Private Cloud (VPC) —Vous devez sélectionner l' Amazon VPC emplacement de votre Amazon FSx (Windows). Vous incluez le sous-réseau VPC et les groupes de sécurité. Voir Configuration d'un Amazon VPC.

e. IAM rôle —Choisissez un IAM rôle existant ou créez-en un nouveau IAM pour accéder aux informations d'identification de votre référentiel et indexer le contenu.

   > **Note**
   >
   > IAM les rôles utilisés pour les index ne peuvent pas être utilisés pour les sources de données. Si vous ne savez pas si un rôle existant est utilisé pour un index ou une FAQ, choisissez Créer un nouveau rôle pour éviter les erreurs.

f. Choisissez Suivant.

7. Sur la page Configurer les paramètres de synchronisation, entrez les informations suivantes :

   a. Étendue de synchronisation, modèles Regex : ajoutez des modèles d'expressions régulières pour inclure ou exclure certains fichiers.

   b. Mode de synchronisation : choisissez la manière dont vous souhaitez mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation.

- Synchronisation complète : Indexez tout le contenu fraîchement, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
- Synchronisation nouvelle, modifiée ou supprimée : Indexez uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

c. Calendrier d'exécution de synchronisation : pour Fréquence, choisissez la fréquence à laquelle vous souhaitez synchroniser le contenu de votre source de données et mettre à jour votre index.

d. Choisissez Suivant.

8. Sur la page Définir les mappages de champs, entrez les informations suivantes :

   a. Sélectionnez parmi les champs par défaut Amazon Kendra générés de vos fichiers que vous souhaitez mapper à votre index. Pour ajouter des champs de source de données personnalisés, créez un nom de champ d'index à mapper et le type de données du champ.

   b. Choisissez Suivant.

9. Sur la page Réviser et créer, vérifiez que les informations saisies sont correctes, puis sélectionnez Ajouter une source de données. Vous pouvez également choisir de modifier vos informations depuis cette page. Votre source de données apparaîtra sur la page Sources de données une fois que la source de données aura été ajoutée avec succès.

API

Pour vous connecter Amazon Kendra à votre système de fichiers Amazon FSx (Windows)

Vous devez spécifier un JSON du schéma de source de données à l'aide du TemplateConfigurationAPI. Vous devez fournir les informations suivantes :

- Source de données —Spécifiez le type de source de données comme `FSX` lorsque vous utilisez TemplateConfigurationSchéma JSON. Spécifiez également la source de données comme `TEMPLATE` lorsque vous appelez le CreateDataSourceAPI.

- ID du système de fichiers : identifiant du système de fichiers Amazon FSx (Windows). Vous trouverez l'ID de votre système de fichiers sur le tableau de bord des systèmes de fichiers de la console Amazon FSx (Windows).
- Type de système de fichiers —Spécifiez le type de système de fichiers en tant que`WINDOWS`.
- Virtual Private Cloud (VPC) **`VpcConfiguration`** —Spécifiez le moment de votre appel. `CreateDataSource` Pour de plus amples informations, veuillez consulter Configuration Amazon Kendra pour utiliser un Amazon VPC.

  > **Note**
  >
  > Vous devez sélectionner l' Amazon VPC emplacement de votre Amazon FSx (Windows). Vous incluez le sous-réseau VPC et les groupes de sécurité.

- Mode de synchronisation : spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation. Vous pouvez choisir entre :
  - `FORCED_FULL_CRAWL`pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
  - `FULL_CRAWL`pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
- Identity Crawler : spécifiez s'il faut activer l'explorateur Amazon Kendra d'identité. Le robot d'exploration d'identité utilise les informations de la liste de contrôle d'accès (ACL) de vos documents pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Si vous disposez d'une ACL pour vos documents et que vous choisissez de l'utiliser, vous pouvez également choisir d'activer le robot d'exploration Amazon Kendra d'identité pour configurer le filtrage des résultats de recherche par contexte utilisateur. Sinon, si le robot d'identification est désactivé, tous les documents peuvent être consultés publiquement. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'exploration d'identité est désactivé, vous pouvez également utiliser l'PutPrincipalMappingAPI pour télécharger les informations d'accès des utilisateurs et des groupes afin de filtrer le contexte utilisateur.

- Nom de ressource Amazon (ARN) secret : indiquez le nom de ressource Amazon (ARN) d'un Secrets Manager secret contenant les informations d'authentification de votre compte Amazon FSx (Windows). Le secret est stocké dans une structure JSON avec les clés suivantes :

```
{
    "username": "user@corp.example.com",
    "password": "password"
}
```

- IAM role —Spécifiez `CreateDataSource` à quel `RoleArn` moment vous appelez pour fournir à un IAM rôle les autorisations d'accéder à votre Secrets Manager secret et pour appeler le public requis APIs pour le connecteur Amazon FSx (Windows) et Amazon Kendra. Pour plus d'informations, consultez la section IAM Rôles pour les sources de données Amazon FSx (Windows).

Vous pouvez également ajouter les fonctionnalités optionnelles suivantes :

- Filtres d'inclusion et d'exclusion : spécifiez si vous souhaitez inclure ou exclure certains fichiers.

> **Note**
>
> La plupart des sources de données utilisent des modèles d'expressions régulières, qui sont des modèles d'inclusion ou d'exclusion appelés filtres. Si vous spécifiez un filtre d'inclusion, seul le contenu correspondant au filtre d'inclusion est indexé. Tout document qui ne correspond pas au filtre d'inclusion n'est pas indexé. Si vous spécifiez un filtre d'inclusion et d'exclusion, les documents correspondant au filtre d'exclusion ne sont pas indexés, même s'ils correspondent au filtre d'inclusion.

- Liste de contrôle d'accès (ACL) : indiquez si vous souhaitez analyser les informations ACL de vos documents, si vous disposez d'une ACL et souhaitez l'utiliser pour le contrôle d'accès. L'ACL indique les documents auxquels les utilisateurs et les groupes peuvent accéder. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.

> **Note**
>
> Pour tester le filtrage du contexte utilisateur sur un utilisateur, vous devez inclure le nom de domaine DNS dans le nom d'utilisateur lorsque vous émettez la requête. Vous devez disposer des autorisations administratives du domaine Active Directory. Vous pouvez également tester le filtrage du contexte utilisateur sur le nom d'un groupe.

- Mappages de champs : choisissez de mapper les champs de votre source de données Amazon FSx (Windows) à vos champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.

> **Note**
>
> Le champ du corps du document ou l'équivalent du corps du document pour vos documents est requis pour Amazon Kendra effectuer une recherche dans vos documents. Vous devez associer le nom du champ du corps du document dans votre source de données au nom du champ d'index`_document_body`. Tous les autres champs sont facultatifs.

Pour obtenir la liste des autres clés JSON importantes à configurer, consultez le schéma de modèle Amazon FSx (Windows).

En savoir plus

Pour en savoir plus sur l'intégration Amazon Kendra à votre source de données Amazon FSx (Windows), consultez :

- Recherchez en toute sécurité des données non structurées sur les systèmes de fichiers Windows avec le Amazon Kendra connecteur pour Amazon FSx (Windows) pour Windows File Server.

Amazon FSx (NetApp ONTAP)

Amazon FSx (NetApp ONTAP) est un système de serveur de fichiers entièrement géré basé sur le cloud qui offre des fonctionnalités de stockage partagé. Si vous êtes un utilisateur Amazon FSx

(NetApp ONTAP), vous pouvez l'utiliser Amazon Kendra pour indexer votre source de données Amazon FSx (NetApp ONTAP).

Vous pouvez vous connecter Amazon Kendra à votre source de données Amazon FSx (NetApp ONTAP) à l'aide de la Amazon Kendra console ou du TemplateConfigurationAPI.

Pour résoudre les problèmes liés à votre connecteur de source de données Amazon Kendra Amazon FSx (NetApp ONTAP), consultezDépannage des sources de données.

Rubriques

- Fonctionnalités prises en charge
- Prérequis
- Instructions de connexion

Fonctionnalités prises en charge

Amazon Kendra Amazon FSx Le connecteur de source de données (NetApp ONTAP) prend en charge les fonctionnalités suivantes :

- Mappages de champs
- Contrôle d'accès utilisateur
- Filtres d'inclusion et d'exclusion
- Synchronisation complète et incrémentielle du contenu
- Cloud privé virtuel (VPC)

Prérequis

Avant de pouvoir Amazon Kendra indexer votre source de données Amazon FSx (NetApp ONTAP), vérifiez les détails de votre source de données Amazon FSx (NetApp ONTAP) et. Comptes AWS

Pour Amazon FSx (NetApp ONTAP), assurez-vous d'avoir :

- Configuration Amazon FSx (NetApp ONTAP) avec autorisations de lecture et de montage.
- Vous avez noté l'identifiant de votre système de fichiers. Vous trouverez l'ID de votre système de fichiers sur le tableau de bord des systèmes de fichiers de la console Amazon FSx (NetApp ONTAP).

- Notez l'ID de la machine virtuelle de stockage (SVM) utilisé avec votre système de fichiers. Vous pouvez trouver votre identifiant de SVM en accédant au tableau de bord des systèmes de fichiers de la console Amazon FSx (NetApp ONTAP), en sélectionnant l'identifiant de votre système de fichiers, puis en sélectionnant Machines virtuelles de stockage.
- Configuré un cloud privé virtuel en utilisant l' Amazon VPC emplacement de votre système de fichiers Amazon FSx (NetApp ONTAP).
- Vous avez noté vos informations d'authentification Amazon FSx (NetApp ONTAP) pour un Active Directory compte utilisateur. Cela inclut votre nom d'utilisateur Active Directory avec votre nom de domaine DNS (par exemple, user@corp.example.com) et votre mot de passe. Si vous utilisez le protocole NFS (Network File System) pour votre système de fichiers Amazon FSx (NetApp ONTAP), les informations d'authentification incluent un identifiant gauche, un identifiant droit et une clé pré-partagée.

  > **Note**
  >
  > Utilisez uniquement les informations d'identification nécessaires au fonctionnement du connecteur. N'utilisez pas d'informations d'identification privilégiées telles que celles d'administrateur de domaine.

  > **Note**
  >
  > Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

- Il est vérifié que chaque document est unique dans Amazon FSx (NetApp ONTAP) et dans les autres sources de données que vous prévoyez d'utiliser pour le même index. Chaque source de données que vous souhaitez utiliser pour un index ne doit pas contenir le même document dans toutes les sources de données. IDs Les documents sont globaux par rapport à un index et doivent être uniques par index.

Dans votre Compte AWS, assurez-vous d'avoir :

- Création d'un Amazon Kendra index et, si vous utilisez l'API, notez l'ID de l'index.

- Vous avez créé un IAM rôle pour votre source de données et, si vous utilisez l'API, notez l'ARN du IAM rôle.

> **Note**
>
> Si vous modifiez votre type d'authentification et vos informations d'identification, vous devez mettre à jour votre IAM rôle pour accéder au bon identifiant AWS Secrets Manager secret.

- Stockez vos informations d'authentification Amazon FSx (NetApp ONTAP) dans un AWS Secrets Manager secret et, si vous utilisez l'API, notez l'ARN du secret.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

Si vous n'avez pas de IAM rôle ou de secret existant, vous pouvez utiliser la console pour créer un nouveau IAM rôle et un nouveau Secrets Manager secret lorsque vous connectez votre source de données Amazon FSx (NetApp ONTAP) à Amazon Kendra. Si vous utilisez l'API, vous devez fournir l'ARN d'un IAM rôle et d'un Secrets Manager secret existants, ainsi qu'un identifiant d'index.

Instructions de connexion

Pour vous connecter Amazon Kendra à votre source de données Amazon FSx (NetApp ONTAP), vous devez fournir les informations nécessaires sur votre source de données Amazon FSx (NetApp ONTAP) afin de Amazon Kendra pouvoir accéder à vos données. Si vous n'avez pas encore configuré Amazon FSx (NetApp ONTAP) pour Amazon Kendra, consultezPrérequis.

Console

Pour vous connecter Amazon Kendra à votre système de fichiers Amazon FSx (NetApp ONTAP)

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la.

2. Dans le volet de navigation de gauche, choisissez Index, puis choisissez l'index que vous souhaitez utiliser dans la liste des index.

> **Note**
>
> Vous pouvez choisir de configurer ou de modifier vos paramètres de contrôle d'accès utilisateur dans les paramètres de l'index.

3. Sur la page de démarrage, choisissez Ajouter une source de données.
4. Sur la page Ajouter une source de données, choisissez le connecteur Amazon FSx (NetApp ONTAP), puis choisissez Ajouter un connecteur. Si vous utilisez la version 2 (le cas échéant), choisissez le connecteur Amazon FSx (NetApp ONTAP) avec la balise « V2.0 ».
5. Sur la page Spécifier les détails de la source de données, entrez les informations suivantes :

   a. Dans Nom et description, pour Nom de la source de données : entrez le nom de votre source de données. Vous pouvez inclure des traits d'union, mais pas des espaces.

   b. (Facultatif) Description : entrez une description facultative pour votre source de données.

   c. Dans la langue par défaut : choisissez une langue pour filtrer vos documents pour l'index. Sauf indication contraire, la langue par défaut est l'anglais. La langue spécifiée dans les métadonnées du document remplace la langue sélectionnée.

   d. Dans Balises, pour Ajouter une nouvelle balise : incluez des balises facultatives pour rechercher et filtrer vos ressources ou suivre vos AWS coûts.

   e. Choisissez Suivant.

6. Sur la page Définir l'accès et la sécurité, entrez les informations suivantes :

   a. Source —Fournissez les informations relatives à votre système de fichiers.

      - Protocole du système de fichiers : choisissez le protocole de votre système de fichiers Amazon FSx (NetApp ONTAP). Vous pouvez choisir le protocole CIFS (Common Internet File System) ou le protocole NFS (Network File System) pour Linux.
      - Amazon FSx ID du système de fichiers (NetApp ONTAP) —Sélectionnez dans le menu déroulant votre identifiant de système de fichiers existant, extrait de (ONTAP). Amazon FSx NetApp Vous pouvez également créer un système de fichiers Amazon FSx (NetApp ONTAP). Vous trouverez l'ID de votre système de fichiers sur le tableau de bord des systèmes de fichiers de la console Amazon FSx (NetApp ONTAP).

- Identifiant SVMAmazon FSx (NetApp ONTAP) pour NetApp ONTAP uniquement) — Indiquez l'ID de machine virtuelle de stockage (SVM) de votre Amazon FSx (NetApp ONTAP) NetApp ONTAP. Vous pouvez trouver votre identifiant de SVM en accédant au tableau de bord des systèmes de fichiers de la console Amazon FSx (NetApp ONTAP), en sélectionnant l'identifiant de votre système de fichiers et en sélectionnant Machines virtuelles de stockage.

b. Autorisation : activez ou désactivez les informations de la liste de contrôle d'accès (ACL) pour vos documents, si vous disposez d'une ACL et que vous souhaitez l'utiliser pour le contrôle d'accès. L'ACL indique les documents auxquels les utilisateurs et les groupes peuvent accéder. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.

c. Authentification : choisissez un AWS Secrets Manager secret existant ou créez-en un nouveau pour stocker les informations d'identification de votre système de fichiers. Si vous choisissez de créer un nouveau secret, une fenêtre AWS Secrets Manager secrète s'ouvre.

   Fournissez un secret qui stocke vos informations d'authentification, à savoir votre nom d'utilisateur et votre mot de passe. Le nom d'utilisateur doit inclure votre nom de domaine DNS. Par exemple, user@corp.example.com.

   Si vous utilisez le protocole NFS pour votre système de fichiers Amazon FSx (NetApp ONTAP), fournissez un secret qui stocke vos informations d'authentification (ID gauche, ID droit et clé pré-partagée).

   Enregistrez et ajoutez votre secret.

d. Virtual Private Cloud (VPC) —Vous devez sélectionner l' Amazon VPC endroit où réside votre Amazon FSx (ONTAP). NetApp Vous incluez le sous-réseau VPC et les groupes de sécurité. Voir Configuration d'un Amazon VPC.

e. IAM rôle —Choisissez un IAM rôle existant ou créez-en un nouveau IAM pour accéder aux informations d'identification de votre référentiel et indexer le contenu.

> **Note**
>
> IAM les rôles utilisés pour les index ne peuvent pas être utilisés pour les sources de données. Si vous ne savez pas si un rôle existant est utilisé pour un index ou une FAQ, choisissez Créer un nouveau rôle pour éviter les erreurs.

f. Choisissez Suivant.

7. Sur la page Configurer les paramètres de synchronisation, entrez les informations suivantes :

   a. Étendue de synchronisation, modèles Regex : ajoutez des modèles d'expressions régulières pour inclure ou exclure certains fichiers.

   b. Mode de synchronisation : choisissez la manière dont vous souhaitez mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation.

      - Synchronisation complète : Indexez tout le contenu fraîchement, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
      - Synchronisation nouvelle, modifiée ou supprimée : Indexez uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

   c. Calendrier d'exécution de synchronisation : pour Fréquence, choisissez la fréquence à laquelle vous souhaitez synchroniser le contenu de votre source de données et mettre à jour votre index.

   d. Choisissez Suivant.

8. Sur la page Définir les mappages de champs, entrez les informations suivantes :

   a. Sélectionnez parmi les champs par défaut Amazon Kendra générés de vos fichiers que vous souhaitez mapper à votre index. Pour ajouter des champs de source de données personnalisés, créez un nom de champ d'index à mapper et le type de données du champ.

b. Choisissez Suivant.

9. Sur la page Réviser et créer, vérifiez que les informations saisies sont correctes, puis sélectionnez Ajouter une source de données. Vous pouvez également choisir de modifier vos informations depuis cette page. Votre source de données apparaîtra sur la page Sources de données une fois qu'elle aura été ajoutée avec succès.

API

Pour vous connecter Amazon Kendra à votre système de fichiers Amazon FSx (NetApp ONTAP)

Vous devez spécifier un JSON du schéma de source de données à l'aide du TemplateConfigurationAPI. Vous devez fournir les informations suivantes :

- Source de données —Spécifiez le type de source de données `FSXONTAP` lorsque vous utilisez le TemplateConfigurationSchéma JSON. Spécifiez également la source de données comme `TEMPLATE` lorsque vous appelez le CreateDataSourceAPI.
- ID du système de fichiers : identifiant du système de fichiers Amazon FSx (NetApp ONTAP). Vous trouverez l'ID de votre système de fichiers sur le tableau de bord des systèmes de fichiers de la console Amazon FSx (NetApp ONTAP).
- ID de la SVM : identifiant de la machine virtuelle de stockage (SVM) utilisé avec votre système de fichiers. Vous pouvez trouver votre identifiant de SVM en accédant au tableau de bord des systèmes de fichiers de la console Amazon FSx (NetApp ONTAP), en sélectionnant l'identifiant de votre système de fichiers, puis en sélectionnant Machines virtuelles de stockage.
- Type de protocole : indiquez si vous utilisez le protocole CIFS (Common Internet File System) ou le protocole NFS (Network File System) pour Linux.
- Type de système de fichiers : spécifiez le type de système de fichiers comme l'un ou l'autre`FSXONTAP`.
- Virtual Private Cloud (VPC) **`VpcConfiguration`** —Spécifiez le moment de votre appel. `CreateDataSource` Pour de plus amples informations, veuillez consulter Configuration Amazon Kendra pour utiliser un Amazon VPC.

> **Note**
>
> Vous devez sélectionner le Amazon VPC lieu de résidence de votre Amazon FSx (NetApp ONTAP). Vous incluez le sous-réseau VPC et les groupes de sécurité.

- Nom de ressource Amazon (ARN) secret : indiquez le nom de ressource Amazon (ARN) d'un Secrets Manager secret contenant les informations d'authentification de votre compte Amazon FSx (NetApp ONTAP). Le secret est stocké dans une structure JSON avec les clés suivantes :

```
{
    "username": "user@corp.example.com",
    "password": "password"
}
```

Si vous utilisez le protocole NFS pour votre système de fichiers Amazon FSx (NetApp ONTAP), le secret est stocké dans une structure JSON avec les clés suivantes :

```
{
    "leftId": "left ID",
    "rightId": "right ID",
    "preSharedKey": "pre-shared key"
}
```

- IAM role —Spécifiez à quel `RoleArn` moment vous appelez `CreateDataSource` pour fournir à un IAM rôle les autorisations d'accéder à votre Secrets Manager secret et d'appeler le public requis APIs pour le connecteur Amazon FSx (NetApp ONTAP) et. Amazon Kendra Pour plus d'informations, consultez la section IAM Rôles pour les sources de données Amazon FSx (NetApp ONTAP).

Vous pouvez également ajouter les fonctionnalités optionnelles suivantes :

- Mode de synchronisation : spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation. Vous pouvez choisir entre :
  - `FORCED_FULL_CRAWL`pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
  - `FULL_CRAWL`pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

- Filtres d'inclusion et d'exclusion : indiquez si vous souhaitez inclure ou exclure certains fichiers.

  > **Note**
  >
  > La plupart des sources de données utilisent des modèles d'expressions régulières, qui sont des modèles d'inclusion ou d'exclusion appelés filtres. Si vous spécifiez un filtre d'inclusion, seul le contenu correspondant au filtre d'inclusion est indexé. Tout document qui ne correspond pas au filtre d'inclusion n'est pas indexé. Si vous spécifiez un filtre d'inclusion et d'exclusion, les documents correspondant au filtre d'exclusion ne sont pas indexés, même s'ils correspondent au filtre d'inclusion.

- Liste de contrôle d'accès (ACL) : indiquez si vous souhaitez analyser les informations ACL de vos documents, si vous disposez d'une ACL et souhaitez l'utiliser pour le contrôle d'accès. L'ACL indique les documents auxquels les utilisateurs et les groupes peuvent accéder. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.

  > **Note**
  >
  > Pour tester le filtrage du contexte utilisateur sur un utilisateur, vous devez inclure le nom de domaine DNS dans le nom d'utilisateur lorsque vous émettez la requête. Vous devez disposer des autorisations administratives du domaine Active Directory. Vous pouvez également tester le filtrage du contexte utilisateur sur le nom d'un groupe.

- Mappages de champs : choisissez de mapper les champs de votre source de données Amazon FSx (NetApp ONTAP) à vos Amazon Kendra champs d'index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.

  > **Note**
  >
  > Le champ du corps du document ou l'équivalent du corps du document pour vos documents est requis pour Amazon Kendra effectuer une recherche dans vos documents. Vous devez associer le nom du champ du corps du document dans votre source de données au nom du champ d'index`_document_body`. Tous les autres champs sont facultatifs.

Pour obtenir la liste des autres clés JSON importantes à configurer, consultez le schéma du modèle Amazon FSx (NetApp ONTAP).

Amazon RDS/Aurora

Vous pouvez indexer des documents stockés dans une base de données à l'aide d'une source de données de base de données. Après avoir fourni les informations de connexion pour la base de données, Amazon Kendra connecte et indexe les documents.

Amazon Kendra prend en charge les bases de données suivantes :

- Amazon Aurora MySQL
- Amazon Aurora PostgreSQL
- Amazon RDS pour MySQL
- Amazon RDS pour PostgreSQL

> **Note**
>
> Les bases de données Aurora sans serveur ne sont pas prises en charge.

> **Important**
>
> La dépréciation de ce connecteur Amazon RDS/Aurora est prévue pour fin 2023. Amazon Kendra prend désormais en charge les nouveaux connecteurs de source de données de base de données. Pour une meilleure expérience, nous vous recommandons de choisir l'un des nouveaux connecteurs suivants pour votre cas d'utilisation :
>
> - Aurora (MySQL)
> - Aurora (PostgreSQL)
> - Amazon RDS (MySQL)
> - Amazon RDS (Microsoft SQL Server)
> - Amazon RDS (Oracle)
> - Amazon RDS (PostgreSQL)
> - IBM DB2

- Microsoft SQL Server
- MySQL
- Base de données Oracle
- PostgreSQL

Vous pouvez vous connecter Amazon Kendra à la source de données de votre base de données à l'aide de la Amazon Kendra console et de l'DatabaseConfigurationAPI.

Pour résoudre les problèmes liés à votre connecteur de source de données de Amazon Kendra base de données, consultezDépannage des sources de données.

Rubriques

- Fonctionnalités prises en charge
- Prérequis
- Instructions de connexion

Fonctionnalités prises en charge

Amazon Kendra le connecteur de source de données de base de données prend en charge les fonctionnalités suivantes :

- Mappages de champs
- Filtrage du contexte utilisateur
- Cloud privé virtuel (VPC)

Prérequis

Avant de pouvoir Amazon Kendra indexer la source de données de votre base de données, apportez ces modifications à votre base de données et à vos AWS comptes.

Dans votre base de données, assurez-vous d'avoir :

- Notez vos informations d'authentification de base, à savoir le nom d'utilisateur et le mot de passe de votre base de données.

- Vous avez copié le nom d'hôte, le numéro de port, l'adresse de l'hôte, le nom de la base de données et le nom de la table de données contenant les données du document. Pour PostgreSQL, la table de données doit être une table publique ou un schéma public.

  > **Note**
  >
  > L'hôte et le port indiquent Amazon Kendra où se trouve le serveur de base de données sur Internet. Le nom de la base de données et le nom de la table indiquent Amazon Kendra où se trouvent les données du document sur le serveur de base de données.

- Les noms des colonnes de la table de données contenant les données du document ont été copiés. Vous devez inclure l'ID du document, le corps du document, les colonnes pour détecter si un document a changé (par exemple, dernière colonne mise à jour) et les colonnes facultatives de la table de données qui correspondent à des champs d'index personnalisés. Vous pouvez également associer n'importe quel nom de champ Amazon Kendra réservé à une colonne de table.
- Vous avez copié les informations relatives au type de moteur de base de données, par exemple si vous l'utilisez Amazon RDS pour MySQL ou un autre type.
- Il est vérifié que chaque document est unique dans la base de données et dans les autres sources de données que vous prévoyez d'utiliser pour le même index. Chaque source de données que vous souhaitez utiliser pour un index ne doit pas contenir le même document dans toutes les sources de données. IDs Les documents sont globaux par rapport à un index et doivent être uniques par index.

Dans votre Compte AWS, assurez-vous d'avoir :

- Création d'un Amazon Kendra index et, si vous utilisez l'API, notez l'ID de l'index.
- Vous avez créé un IAM rôle pour votre source de données et, si vous utilisez l'API, notez l'ARN du IAM rôle.

  > **Note**
  >
  > Si vous modifiez votre type d'authentification et vos informations d'identification, vous devez mettre à jour votre IAM rôle pour accéder au bon identifiant AWS Secrets Manager secret.

- Stockez les informations d'authentification de votre base de données dans un AWS Secrets Manager secret et, si vous utilisez l'API, notez l'ARN du secret.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

Si vous n'avez pas de IAM rôle ou de secret existant, vous pouvez utiliser la console pour créer un nouveau IAM rôle et un nouveau Secrets Manager secret lorsque vous connectez votre source de données de base de données à Amazon Kendra. Si vous utilisez l'API, vous devez fournir l'ARN d'un IAM rôle et d'un Secrets Manager secret existants, ainsi qu'un identifiant d'index.

Instructions de connexion

Pour vous connecter Amazon Kendra à votre source de données de base de données, vous devez fournir les informations nécessaires sur votre source de données de base de données afin de Amazon Kendra pouvoir accéder à vos données. Si vous n'avez pas encore configuré la base de données pour Amazon Kendra, consultezPrérequis.

Console

Pour se connecter Amazon Kendra à une base de données

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la.
2. Dans le volet de navigation de gauche, choisissez Index, puis choisissez l'index que vous souhaitez utiliser dans la liste des index.

   > **Note**
   >
   > Vous pouvez choisir de configurer ou de modifier vos paramètres de contrôle d'accès utilisateur dans les paramètres de l'index.

3. Sur la page de démarrage, choisissez Ajouter une source de données.
4. Sur la page Ajouter une source de données, choisissez le connecteur de base de données, puis sélectionnez Ajouter un connecteur. Si vous utilisez la version 2 (le cas échéant), choisissez le connecteur de base de données avec la balise « V2.0 ».

5. Sur la page Spécifier les détails de la source de données, entrez les informations suivantes :

   a. Dans Nom et description, pour Nom de la source de données : entrez le nom de votre source de données. Vous pouvez inclure des tirets, mais pas des espaces.

   b. (Facultatif) Description : entrez une description facultative pour votre source de données.

   c. Dans la langue par défaut : choisissez une langue pour filtrer vos documents pour l'index. Sauf indication contraire, la langue par défaut est l'anglais. La langue spécifiée dans les métadonnées du document remplace la langue sélectionnée.

   d. Dans Balises, pour Ajouter une nouvelle balise : incluez des balises facultatives pour rechercher et filtrer vos ressources ou suivre vos AWS coûts.

   e. Choisissez Suivant.

6. Sur la page Définir l'accès et la sécurité, entrez les informations suivantes :

   a. Point de terminaison : nom d'hôte DNS, IPv4 adresse ou IPv6 adresse.

   b. Port : numéro de port.

   c. Base de données : nom de la base de données.

   d. Nom de la table —Nom de la table.

   e. Pour Type d'authentification, choisissez entre Existant et Nouveau pour stocker les informations d'authentification de votre base de données. Si vous choisissez de créer un nouveau secret, une fenêtre AWS Secrets Manager secrète s'ouvre.

      - Entrez les informations suivantes dans la fenêtre Créer un AWS Secrets Manager secret :

        A. Nom secret : le nom de votre secret. Le préfixe « AmazonKendra -database-» est automatiquement ajouté à votre nom secret.

        B. Pour le nom d'utilisateur et le mot de passe : entrez les valeurs d'authentification de votre compte de base de données.

        C. Choisissez Enregistrer l'authentification.

   f. Virtual Private Cloud (VPC) —Vous pouvez choisir d'utiliser un VPC. Dans ce cas, vous devez ajouter des sous-réseaux et des groupes de sécurité VPC.

> **Note**
>
> Vous devez utiliser un sous-réseau privé. Si votre instance RDS se trouve dans un sous-réseau public de votre VPC, vous pouvez créer un sous-réseau privé

doté d'un accès sortant à une passerelle NAT dans le sous-réseau public. Les sous-réseaux fournis dans la configuration VPC doivent se trouver dans l'ouest des États-Unis (Oregon), dans l'est des États-Unis (Virginie du Nord) ou dans l'UE (Irlande).

g. IAM rôle —Choisissez un IAM rôle existant ou créez-en un nouveau IAM pour accéder aux informations d'identification de votre référentiel et indexer le contenu.

> **Note**
>
> IAM les rôles utilisés pour les index ne peuvent pas être utilisés pour les sources de données. Si vous ne savez pas si un rôle existant est utilisé pour un index ou une FAQ, choisissez Créer un nouveau rôle pour éviter les erreurs.

h. Choisissez Suivant.

7. Sur la page Configurer les paramètres de synchronisation, entrez les informations suivantes :

   a. Choisissez entre Aurora MySQL, MySQL, Aurora PostgreSQL et PostgreSQL en fonction de votre cas d'utilisation.

   b. Placer les identificateurs SQL entre guillemets : sélectionnez cette option pour placer les identificateurs SQL entre guillemets doubles. Par exemple, « ColumnName ».

   c. Colonne ACL et colonnes de détection des modifications : configurez les colonnes Amazon Kendra utilisées pour la détection des modifications (par exemple, dernière colonne mise à jour) et votre liste de contrôle d'accès.

   d. Dans Synchroniser le calendrier d'exécution, pour Fréquence : choisissez la fréquence de synchronisation avec votre source de données. Amazon Kendra

   e. Choisissez Suivant.

8. Sur la page Définir les mappages de champs, entrez les informations suivantes :

   a. Amazon Kendra mappages de champs par défaut : sélectionnez parmi les champs de source de données par défaut Amazon Kendra générés que vous souhaitez mapper à votre index. Vous devez ajouter les valeurs des colonnes de base de données pour `document_id` et `document_body`

   b. Mappages de champs personnalisés : pour ajouter des champs de source de données personnalisés afin de créer un nom de champ d'index à mapper et le type de données du champ.

c. Choisissez Suivant.

9. Sur la page Réviser et créer, vérifiez que les informations saisies sont correctes, puis sélectionnez Ajouter une source de données. Vous pouvez également choisir de modifier vos informations depuis cette page. Votre source de données apparaîtra sur la page Sources de données une fois que la source de données aura été ajoutée avec succès.

API

Pour se connecter Amazon Kendra à une base de données

Vous devez spécifier l'DatabaseConfigurationAPI suivante :

- ColumnConfiguration—Informations sur l'endroit où l'index doit obtenir les informations du document à partir de la base de données. Pour en savoir plus, consultez ColumnConfiguration. Vous devez spécifier les champs `DocumentDataColumnName` (corps du document ou texte principal) et `DocumentIdColumnName` `ChangeDetectingColumn` (par exemple, dernière colonne mise à jour). La colonne mappée au `DocumentIdColumnName` champ doit être une colonne entière. L'exemple suivant montre une configuration de colonne simple pour une source de données de base de données :

```
"ColumnConfiguration": {
    "ChangeDetectingColumns": [
        "LastUpdateDate",
        "LastUpdateTime"
    ],
    "DocumentDataColumnName": "TextColumn",
    "DocumentIdColumnName": "IdentifierColumn",
    "DocoumentTitleColumnName": "TitleColumn",
    "FieldMappings": [
        {
            "DataSourceFieldName": "AbstractColumn",
            "IndexFieldName": "Abstract"
        }
    ]
}
```

- ConnectionConfiguration: informations de configuration requises pour se connecter à une base de données. Pour en savoir plus, consultez ConnectionConfiguration.
- DatabaseEngineType: type de moteur de base de données qui exécute la base de données. Le `DatabaseHost` champ pour `ConnectionConfiguration` doit être le point de terminaison

de l'instance Amazon Relational Database Service (Amazon RDS) de la base de données. N'utilisez pas le point de terminaison du cluster.

- Nom de ressource Amazon (ARN) secret : indiquez le nom de ressource Amazon (ARN) d'un Secrets Manager secret contenant les informations d'authentification de votre compte de base de données. Le secret est stocké dans une structure JSON avec les clés suivantes :

```
{
    "username": "user name",
    "password": "password"
}
```

L'exemple suivant montre une configuration de base de données, y compris l'ARN secret.

```
"DatabaseConfiguration": {
"ConnectionConfiguration": {
"DatabaseHost": "host.subdomain.domain.tld",
        "DatabaseName": "DocumentDatabase",
        "DatabasePort": 3306,
        "SecretArn": "arn:aws:secretmanager:region:account ID:secret/secret name",
        "TableName": "DocumentTable"
    }
}
```

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

- IAM role —Spécifiez à `RoleArn` quel moment vous appelez `CreateDataSource` pour fournir à un IAM rôle les autorisations d'accéder à votre Secrets Manager secret et pour appeler le public requis APIs pour le connecteur de base de données et Amazon Kendra. Pour plus d'informations, consultez la section IAM Rôles des sources de données de base de données.

Vous pouvez également ajouter les fonctionnalités optionnelles suivantes :

- Virtual Private Cloud (VPC) —Spécifiez dans le `VpcConfiguration` cadre de la configuration de la source de données. Consultez la section Configuration Amazon Kendra pour utiliser un VPC.

  > **Note**
  >
  > Vous ne devez utiliser qu'un sous-réseau privé. Si votre instance RDS se trouve dans un sous-réseau public de votre VPC, vous pouvez créer un sous-réseau privé doté d'un accès sortant à une passerelle NAT dans le sous-réseau public. Les sous-réseaux fournis dans la configuration VPC doivent se trouver dans l'ouest des États-Unis (Oregon), dans l'est des États-Unis (Virginie du Nord) ou dans l'UE (Irlande).

- Mappages de champs : choisissez de mapper les champs de votre source de données de base de données à vos champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.

  > **Note**
  >
  > Le champ du corps du document ou l'équivalent du corps du document pour vos documents est requis pour Amazon Kendra effectuer une recherche dans vos documents. Vous devez associer le nom du champ du corps du document dans votre source de données au nom du champ d'index`_document_body`. Tous les autres champs sont facultatifs.

- Filtrage du contexte utilisateur et contrôle d'accèsAmazon Kendra : analyse la liste de contrôle d'accès (ACL) de vos documents, si vous disposez d'une ACL pour vos documents. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.

Amazon RDS (Microsoft SQL Server)

> **Note**
>
> Amazon RDS Le connecteur (Microsoft SQL Server) reste entièrement pris en charge pour les clients existants jusqu'au 31 mai 2026. Bien que ce connecteur ne soit plus disponible pour les nouveaux utilisateurs, les utilisateurs actuels peuvent continuer à l'utiliser sans

interruption. Nous faisons constamment évoluer notre portefeuille de connecteurs afin de proposer des solutions plus évolutives et personnalisables. Pour les intégrations futures, nous vous recommandons d'explorer le framework de connecteurs personnalisés Amazon Kendra [1], conçu pour prendre en charge un plus large éventail de cas d'utilisation en entreprise avec une flexibilité accrue.

SQL Server est un système de gestion de base de données développé par Microsoft. Amazon RDS for SQL Server facilite la configuration, l'exploitation et le dimensionnement des déploiements de SQL Server dans le cloud. Si vous êtes un utilisateur Amazon RDS (Microsoft SQL Server), vous pouvez l'utiliser Amazon Kendra pour indexer votre source de données Amazon RDS (Microsoft SQL Server). Le connecteur de source de données Amazon Kendra JDBC prend en charge Microsoft SQL Server 2019.

Vous pouvez vous connecter Amazon Kendra à votre source de données Amazon RDS (Microsoft SQL Server) à l'aide de la Amazon Kendra console et de l'TemplateConfigurationAPI.

Pour résoudre les problèmes liés à votre connecteur de source de données Amazon Kendra Amazon RDS (Microsoft SQL Server), consultezDépannage des sources de données.

Rubriques

- Fonctionnalités prises en charge
- Prérequis
- Instructions de connexion
- Remarques

Fonctionnalités prises en charge

- Mappages de champs
- Filtrage du contexte utilisateur
- Filtres d'inclusion/exclusion
- Synchronisation complète et incrémentielle du contenu
- Cloud privé virtuel (VPC)

Prérequis

Avant de pouvoir Amazon Kendra indexer votre source de données Amazon RDS (Microsoft SQL Server), apportez ces modifications à votre Amazon RDS (Microsoft SQL Server) et à vos AWS comptes.

Dans Amazon RDS (Microsoft SQL Server), assurez-vous d'avoir :

- Notez le nom d'utilisateur et le mot de passe de votre base de données.

  > **Important**
  >
  > Il est recommandé de fournir des informations d'identification Amazon Kendra de base de données en lecture seule.

- Vous avez copié l'URL, le port et l'instance de votre hôte de base de données.
- Il est vérifié que chaque document est unique dans Amazon RDS (Microsoft SQL Server) et dans les autres sources de données que vous prévoyez d'utiliser pour le même index. Chaque source de données que vous souhaitez utiliser pour un index ne doit pas contenir le même document dans toutes les sources de données. IDs Les documents sont globaux par rapport à un index et doivent être uniques par index.

Dans votre Compte AWS, assurez-vous d'avoir :

- Création d'un Amazon Kendra index et, si vous utilisez l'API, notez l'ID de l'index.
- Vous avez créé un IAM rôle pour votre source de données et, si vous utilisez l'API, notez l'ARN du IAM rôle.

  > **Note**
  >
  > Si vous modifiez votre type d'authentification et vos informations d'identification, vous devez mettre à jour votre IAM rôle pour accéder au bon identifiant AWS Secrets Manager secret.

- Stockez vos informations d'authentification Amazon RDS (Microsoft SQL Server) dans un AWS Secrets Manager secret et, si vous utilisez l'API, notez l'ARN du secret.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

Si vous n'avez pas de IAM rôle ou de secret existant, vous pouvez utiliser la console pour créer un nouveau IAM rôle et un nouveau Secrets Manager secret lorsque vous connectez votre source de données Amazon RDS (Microsoft SQL Server) à Amazon Kendra. Si vous utilisez l'API, vous devez fournir l'ARN d'un IAM rôle et d'un Secrets Manager secret existants, ainsi qu'un identifiant d'index.

Instructions de connexion

Pour vous connecter Amazon Kendra à votre source de données Amazon RDS (Microsoft SQL Server), vous devez fournir les détails de vos informations d'identification Amazon RDS (Microsoft SQL Server) afin de Amazon Kendra pouvoir accéder à vos données. Si vous n'avez pas encore configuré Amazon RDS (Microsoft SQL Server) pour Amazon Kendra voirPrérequis.

Console

Pour se connecter Amazon Kendra à Amazon RDS (Microsoft SQL Server)

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la.
2. Dans le volet de navigation de gauche, choisissez Index, puis choisissez l'index que vous souhaitez utiliser dans la liste des index.

   > **Note**
   >
   > Vous pouvez choisir de configurer ou de modifier vos paramètres de contrôle d'accès utilisateur dans les paramètres de l'index.

3. Sur la page de démarrage, choisissez Ajouter une source de données.
4. Sur la page Ajouter une source de données, choisissez le connecteur Amazon RDS (Microsoft SQL Server), puis choisissez Ajouter un connecteur. Si vous utilisez la version 2 (le

cas échéant), choisissez le connecteur Amazon RDS (Microsoft SQL Server) avec la balise « V2.0 ».

5. Sur la page Spécifier les détails de la source de données, entrez les informations suivantes :

   a. Dans Nom et description, pour Nom de la source de données : entrez le nom de votre source de données. Vous pouvez inclure des tirets, mais pas des espaces.

   b. (Facultatif) Description : entrez une description facultative pour votre source de données.

   c. Dans la langue par défaut : choisissez une langue pour filtrer vos documents pour l'index. Sauf indication contraire, la langue par défaut est l'anglais. La langue spécifiée dans les métadonnées du document remplace la langue sélectionnée.

   d. Dans Balises, pour Ajouter une nouvelle balise : incluez des balises facultatives pour rechercher et filtrer vos ressources ou suivre vos AWS coûts.

   e. Choisissez Suivant.

6. Sur la page Définir l'accès et la sécurité, entrez les informations suivantes :

   a. Dans Source, entrez les informations suivantes :

   b. Hôte : entrez le nom d'hôte de la base de données.

   c. Port — Entrez le port de base de données.

   d. Instance — Entrez l'instance de base de données.

   e. Activer l'emplacement du certificat SSL : choisissez d'entrer le Amazon S3 chemin d'accès à votre fichier de certificat SSL.

   f. Dans Authentification, entrez les informations suivantes :

      - AWS Secrets Manager secret —Choisissez un secret existant ou créez-en un nouveau Secrets Manager pour stocker vos informations d'authentification Amazon RDS (Microsoft SQL Server). Si vous choisissez de créer un nouveau secret, une fenêtre AWS Secrets Manager secrète s'ouvre.

        A. Entrez les informations suivantes dans la fenêtre Créer un AWS Secrets Manager secret :

           I. Nom secret : le nom de votre secret. Le préfixe « AmazonKendra -Amazon RDS (Microsoft SQL Server) -» est automatiquement ajouté à votre nom secret.

II. Pour le nom d'utilisateur et le mot de passe de la base de données, entrez les valeurs d'identification d'authentification que vous avez copiées depuis votre base de données.

B. Choisissez Enregistrer.

g. Virtual Private Cloud (VPC) —Vous pouvez choisir d'utiliser un VPC. Dans ce cas, vous devez ajouter des sous-réseaux et des groupes de sécurité VPC.

h. IAM rôle —Choisissez un IAM rôle existant ou créez-en un nouveau IAM pour accéder aux informations d'identification de votre référentiel et indexer le contenu.

> **Note**
>
> IAM les rôles utilisés pour les index ne peuvent pas être utilisés pour les sources de données. Si vous ne savez pas si un rôle existant est utilisé pour un index ou une FAQ, choisissez Créer un nouveau rôle pour éviter les erreurs.

i. Choisissez Suivant.

7. Sur la page Configurer les paramètres de synchronisation, entrez les informations suivantes :

a. Dans le champ d'application de la synchronisation, choisissez l'une des options suivantes :

- Requête SQL —Entrez des instructions de requête SQL telles que les opérations SELECT et JOIN. La taille des requêtes SQL doit être inférieure à 32 Ko. Amazon Kendra analysera tout le contenu de la base de données correspondant à votre requête.

> **Note**
>
> Si le nom d'une table contient des caractères spéciaux (non alphanumériques), vous devez utiliser des crochets autour du nom de la table. Par exemple, `select * from [my-database-table]`

- Colonne clé primaire : indiquez la clé primaire pour la table de base de données. Cela identifie une table au sein de votre base de données.
- Colonne de titre : indiquez le nom de la colonne de titre du document dans votre table de base de données.

- Colonne du corps : indiquez le nom de la colonne du corps du document dans votre table de base de données.

b. Dans Configuration supplémentaire — facultatif, choisissez l'une des options suivantes pour synchroniser un contenu spécifique au lieu de synchroniser tous les fichiers :

- Colonnes détectant les modifications : entrez le nom des colonnes qui Amazon Kendra seront utilisées pour détecter les modifications de contenu. Amazon Kendra réindexera le contenu en cas de modification de l'une de ces colonnes.
- IDsColonne utilisateur : entrez le nom de la colonne contenant l'utilisateur devant être autorisé IDs à accéder au contenu.
- Colonne Groupes : entrez le nom de la colonne contenant les groupes autorisés à accéder au contenu.
- URLsColonne source —Entrez le nom de la colonne contenant la source URLs à indexer.
- Colonne horodatage : entrez le nom de la colonne contenant les horodatages. Amazon Kendra utilise les informations d'horodatage pour détecter les modifications apportées à votre contenu et synchroniser uniquement le contenu modifié.
- Colonne fuseaux horaires : entrez le nom de la colonne contenant les fuseaux horaires du contenu à analyser.
- Format d'horodatage : entrez le nom de la colonne contenant les formats d'horodatage à utiliser pour détecter les modifications de contenu et resynchroniser votre contenu.

c. Mode de synchronisation : choisissez la manière dont vous souhaitez mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation.

- Synchronisation complète : Indexez tout le contenu fraîchement, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
- Nouvelle synchronisation modifiée : Indexez uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour

suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

- Synchronisation nouvelle, modifiée ou supprimée : Indexez uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

d. Dans Synchroniser le calendrier d'exécution, pour Fréquence : à quelle fréquence Amazon Kendra sera synchronisée avec votre source de données.

e. Choisissez Suivant.

8. Sur la page Définir les mappages de champs, entrez les informations suivantes :

a. Sélectionnez parmi les champs de source de données par défaut générés (Document IDs, Titres des documents et Source URLs) que vous souhaitez Amazon Kendra mapper à l'index.

b. Ajouter un champ —Pour ajouter des champs de source de données personnalisés afin de créer un nom de champ d'index à mapper et le type de données du champ.

c. Choisissez Suivant.

9. Sur la page Réviser et créer, vérifiez que les informations saisies sont correctes, puis sélectionnez Ajouter une source de données. Vous pouvez également choisir de modifier vos informations depuis cette page. Votre source de données apparaîtra sur la page Sources de données une fois que la source de données aura été ajoutée avec succès.

API

Pour se connecter Amazon Kendra à Amazon RDS (Microsoft SQL Server)

Vous devez spécifier les éléments suivants à l'aide de l'TemplateConfigurationAPI :

- Source de données —Spécifiez le type de source de données tel que `JDBC` lorsque vous utilisez le schéma TemplateConfigurationJSON. Spécifiez également la source de données `TEMPLATE` lorsque vous appelez l'CreateDataSourceAPI.
- Type de base de données —Vous devez spécifier le type de base de données sous `sqlserver` la forme.

- Requête SQL : spécifiez des instructions de requête SQL telles que les opérations SELECT et JOIN. La taille des requêtes SQL doit être inférieure à 32 Ko. Amazon Kendra analysera tout le contenu de la base de données correspondant à votre requête.

  > **Note**
  >
  > Si le nom d'une table contient des caractères spéciaux (non alphanumériques), vous devez utiliser des crochets autour du nom de la table. Par exemple, *`select * from [my-database-table]`*

- Mode de synchronisation : spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation. Vous pouvez choisir entre :
  - `FORCED_FULL_CRAWL`pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
  - `FULL_CRAWL`pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
  - `CHANGE_LOG`pour indexer uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
- Nom de ressource Amazon (ARN) secret : indiquez le nom de ressource Amazon (ARN) d'un Secrets Manager secret contenant les informations d'authentification que vous avez créées dans votre compte Amazon RDS (Microsoft SQL Server). Le secret est stocké dans une structure JSON avec les clés suivantes :

```
{
    "user name": "database user name",
    "password": "password"
}
```

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

- IAM role —Spécifiez `CreateDataSource` à quel `RoleArn` moment vous appelez pour fournir à un IAM rôle les autorisations d'accéder à votre Secrets Manager secret et pour appeler le public requis APIs pour le connecteur Amazon RDS (Microsoft SQL Server) et Amazon Kendra. Pour plus d'informations, consultez la section IAM Rôles pour les sources de données Amazon RDS (Microsoft SQL Server).

Vous pouvez également ajouter les fonctionnalités optionnelles suivantes :

- Virtual Private Cloud (VPC) **`VpcConfiguration`** —Spécifiez le moment de votre appel. `CreateDataSource` Pour de plus amples informations, veuillez consulter Configuration Amazon Kendra pour utiliser un Amazon VPC.
- Filtres d'inclusion et d'exclusion : vous pouvez spécifier si vous souhaitez inclure un contenu spécifique en utilisant l'utilisateur IDs, les groupes, la source URLs, les horodatages et les fuseaux horaires.
- Filtrage du contexte utilisateur et contrôle d'accèsAmazon Kendra : analyse la liste de contrôle d'accès (ACL) de vos documents, si vous disposez d'une ACL pour vos documents. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.
- Mappages de champs : choisissez de mapper les champs de votre source de données (Amazon RDS Microsoft SQL Server) à vos champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.

> **Note**
>
> Le champ du corps du document ou l'équivalent du corps du document pour vos documents est requis pour Amazon Kendra effectuer une recherche dans vos documents. Vous devez associer le nom du champ du corps du document dans votre

source de données au nom du champ d'index_document_body. Tous les autres champs sont facultatifs.

Pour obtenir la liste des autres clés JSON importantes à configurer, consultezAmazon RDS Schéma de modèle (Microsoft SQL Server).

Remarques

- Les lignes de base de données supprimées ne seront pas suivies lors de la Amazon Kendra vérification du contenu mis à jour.
- La taille des noms et des valeurs des champs d'une ligne de votre base de données ne peut pas dépasser 400 Ko.
- Si la source de données de votre base de données contient une grande quantité de données et que vous ne souhaitez pas Amazon Kendra indexer tout le contenu de votre base de données après la première synchronisation, vous pouvez choisir de ne synchroniser que les documents nouveaux, modifiés ou supprimés.
- Il est recommandé de fournir des informations d'identification Amazon Kendra de base de données en lecture seule.
- Il est recommandé d'éviter d'ajouter des tableaux contenant des données sensibles ou des informations personnelles identifiables (PII).

Amazon RDS (MySQL)

Note

Amazon RDS (MySQL)Le connecteur reste entièrement pris en charge pour les clients existants jusqu'au 31 mai 2026. Bien que ce connecteur ne soit plus disponible pour les nouveaux utilisateurs, les utilisateurs actuels peuvent continuer à l'utiliser sans interruption. Nous faisons constamment évoluer notre portefeuille de connecteurs afin de proposer des solutions plus évolutives et personnalisables. Pour les intégrations futures, nous vous recommandons d'explorer le framework de connecteurs personnalisés Amazon Kendra [1], conçu pour prendre en charge un plus large éventail de cas d'utilisation en entreprise avec une flexibilité accrue.

Amazon RDS (Amazon Relational Database Service) est un service Web qui facilite la configuration, l'exploitation et le dimensionnement d'une base de données relationnelle dans AWS le cloud. Si vous êtes un Amazon RDS utilisateur, vous pouvez l'utiliser Amazon Kendra pour indexer votre source de Amazon RDS (MySQL) données. Le connecteur Amazon Kendra de source de données prend en charge les versions Amazon RDS MySql 5.6, 5.7 et 8.0.

Vous pouvez vous connecter Amazon Kendra à votre source de Amazon RDS (MySQL) données à l'aide de la Amazon Kendra console et de l'TemplateConfigurationAPI.

Pour résoudre les problèmes liés à votre connecteur de source de Amazon Kendra Amazon RDS (MySQL) données, consultezDépannage des sources de données.

Rubriques

- Fonctionnalités prises en charge
- Prérequis
- Instructions de connexion
- Remarques

Fonctionnalités prises en charge

- Mappages de champs
- Filtrage du contexte utilisateur
- Filtres d'inclusion/exclusion
- Synchronisation complète et incrémentielle du contenu
- Cloud privé virtuel (VPC)

Prérequis

Avant de pouvoir Amazon Kendra indexer votre source de Amazon RDS (MySQL) données, apportez ces modifications à vos AWS comptes Amazon RDS (MySQL) and.

DansAmazon RDS (MySQL), assurez-vous d'avoir :

- Notez le nom d'utilisateur et le mot de passe de votre base de données.

> **Important**
>
> Il est recommandé de fournir des informations d'identification Amazon Kendra de base de données en lecture seule.

- Vous avez copié l'URL, le port et l'instance de votre hôte de base de données. Vous pouvez trouver ces informations sur la Amazon RDS console.
- Il est vérifié que chaque document est unique dans Amazon RDS (MySQL) et entre les autres sources de données que vous prévoyez d'utiliser pour le même index. Chaque source de données que vous souhaitez utiliser pour un index ne doit pas contenir le même document dans toutes les sources de données. IDs Les documents sont globaux par rapport à un index et doivent être uniques par index.

Dans votre Compte AWS, assurez-vous d'avoir :

- Création d'un Amazon Kendra index et, si vous utilisez l'API, notez l'ID de l'index.
- Vous avez créé un IAM rôle pour votre source de données et, si vous utilisez l'API, notez l'ARN du IAM rôle.

> **Note**
>
> Si vous modifiez votre type d'authentification et vos informations d'identification, vous devez mettre à jour votre IAM rôle pour accéder au bon identifiant AWS Secrets Manager secret.

- Stockez vos informations d'Amazon RDS (MySQL)authentification dans un AWS Secrets Manager secret et, si vous utilisez l'API, notez l'ARN du secret.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

Si vous n'avez pas de IAM rôle ou de secret existant, vous pouvez utiliser la console pour créer un nouveau IAM rôle et un nouveau Secrets Manager secret lorsque vous connectez votre source de Amazon RDS (MySQL) données à Amazon Kendra. Si vous utilisez l'API, vous devez fournir l'ARN d'un IAM rôle et d'un Secrets Manager secret existants, ainsi qu'un identifiant d'index.

Instructions de connexion

Pour vous connecter Amazon Kendra à votre source de Amazon RDS (MySQL) données, vous devez fournir les détails de vos Amazon RDS (MySQL) informations d'identification afin de Amazon Kendra pouvoir accéder à vos données. Si vous n'avez pas encore configuré Amazon RDS (MySQL) pour Amazon Kendra voirPrérequis.

Console

Pour vous connecter Amazon Kendra à Amazon RDS (MySQL)

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la.
2. Dans le volet de navigation de gauche, choisissez Index, puis choisissez l'index que vous souhaitez utiliser dans la liste des index.

   > **Note**
   >
   > Vous pouvez choisir de configurer ou de modifier vos paramètres de contrôle d'accès utilisateur dans les paramètres de l'index.

3. Sur la page de démarrage, choisissez Ajouter une source de données.
4. Sur la page Ajouter une source de données, choisissez Amazon RDS (MySQL)connecteur, puis sélectionnez Ajouter un connecteur. Si vous utilisez la version 2 (le cas échéant), choisissez le Amazon RDS (MySQL)connecteur avec la balise « V2.0 ».
5. Sur la page Spécifier les détails de la source de données, entrez les informations suivantes :

   a. Dans Nom et description, pour Nom de la source de données : entrez le nom de votre source de données. Vous pouvez inclure des tirets, mais pas des espaces.

   b. (Facultatif) Description : entrez une description facultative pour votre source de données.

   c. Dans la langue par défaut : choisissez une langue pour filtrer vos documents pour l'index. Sauf indication contraire, la langue par défaut est l'anglais. La langue spécifiée dans les métadonnées du document remplace la langue sélectionnée.

d. Dans Balises, pour Ajouter une nouvelle balise : incluez des balises facultatives pour rechercher et filtrer vos ressources ou suivre vos AWS coûts.

e. Choisissez Suivant.

6. Sur la page Définir l'accès et la sécurité, entrez les informations suivantes :

a. Dans Source, entrez les informations suivantes :

b. Hôte : entrez l'URL de l'hôte de la base de données, par exemple :`http://`*`instance URL`*`.`*`region`*`.rds.amazonaws.com`.

c. Port — Entrez le port de base de données, par exemple,`5432`.

d. Instance — Entrez l'instance de base de données, par exemple`postgres`.

e. Activer l'emplacement du certificat SSL : choisissez d'entrer le Amazon S3 chemin d'accès à votre fichier de certificat SSL.

f. Dans Authentification, entrez les informations suivantes :

- AWS Secrets Manager secret —Choisissez un secret existant ou créez-en un nouveau Secrets Manager pour stocker vos informations Amazon RDS (MySQL) d'authentification. Si vous choisissez de créer un nouveau secret, une fenêtre AWS Secrets Manager secrète s'ouvre.

  A. Entrez les informations suivantes dans la fenêtre Créer un AWS Secrets Manager secret :

     I. Nom secret : le nom de votre secret. Le préfixe « AmazonKendra - Amazon RDS (MySQL) - » est automatiquement ajouté à votre nom secret.

     II. Pour le nom d'utilisateur et le mot de passe de la base de données, entrez les valeurs d'identification d'authentification que vous avez copiées depuis votre base de données.

  B. Choisissez Enregistrer.

g. Virtual Private Cloud (VPC) —Vous pouvez choisir d'utiliser un VPC. Dans ce cas, vous devez ajouter des sous-réseaux et des groupes de sécurité VPC.

h. IAM rôle —Choisissez un IAM rôle existant ou créez-en un nouveau IAM pour accéder aux informations d'identification de votre référentiel et indexer le contenu.

> **Note**
>
> IAM les rôles utilisés pour les index ne peuvent pas être utilisés pour les sources de données. Si vous ne savez pas si un rôle existant est utilisé pour un index ou une FAQ, choisissez Créer un nouveau rôle pour éviter les erreurs.

i. Choisissez Suivant.

7. Sur la page Configurer les paramètres de synchronisation, entrez les informations suivantes :

   a. Dans le champ d'application de la synchronisation, choisissez l'une des options suivantes :

      - Requête SQL —Entrez des instructions de requête SQL telles que les opérations SELECT et JOIN. Les requêtes SQL doivent être inférieures à 32 Ko Les requêtes SQL doivent être inférieures à 32 Ko et ne pas contenir de points-virgules (;). Amazon Kendra analysera tout le contenu de la base de données correspondant à votre requête.
      - Colonne clé primaire : indiquez la clé primaire pour la table de base de données. Cela identifie une table au sein de votre base de données.
      - Colonne de titre : indiquez le nom de la colonne de titre du document dans votre table de base de données.
      - Colonne du corps : indiquez le nom de la colonne du corps du document dans votre table de base de données.

   b. Dans Configuration supplémentaire (facultatif), choisissez l'une des options suivantes pour synchroniser un contenu spécifique au lieu de synchroniser tous les fichiers :

      - Colonnes détectant les modifications : entrez le nom des colonnes qui Amazon Kendra seront utilisées pour détecter les modifications de contenu. Amazon Kendra réindexera le contenu en cas de modification de l'une de ces colonnes.
      - IDsColonne des utilisateurs : entrez le nom de la colonne contenant l'utilisateur devant être autorisé IDs à accéder au contenu.
      - Colonne Groupes : entrez le nom de la colonne contenant les groupes autorisés à accéder au contenu.
      - URLsColonne source —Entrez le nom de la colonne contenant la source URLs à indexer.

- Colonne horodatage : entrez le nom de la colonne contenant les horodatages. Amazon Kendra utilise les informations d'horodatage pour détecter les modifications apportées à votre contenu et synchroniser uniquement le contenu modifié.
- Colonne fuseaux horaires : entrez le nom de la colonne contenant les fuseaux horaires du contenu à analyser.
- Format d'horodatage : entrez le nom de la colonne contenant les formats d'horodatage à utiliser pour détecter les modifications de contenu et resynchroniser votre contenu.

c. Mode de synchronisation : choisissez la manière dont vous souhaitez mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation.

- Synchronisation complète : Indexez tout le contenu fraîchement, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
- Nouvelle synchronisation modifiée : Indexez uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
- Synchronisation nouvelle, modifiée ou supprimée : Indexez uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

d. Dans Synchroniser le calendrier d'exécution, pour Fréquence : à quelle fréquence Amazon Kendra sera synchronisée avec votre source de données.

e. Choisissez Suivant.

8. Sur la page Définir les mappages de champs, entrez les informations suivantes :

a. Sélectionnez parmi les champs de source de données par défaut générés (Document IDs, Titres des documents et Source URLs) que vous souhaitez Amazon Kendra mapper à l'index.

b. Ajouter un champ —Pour ajouter des champs de source de données personnalisés afin de créer un nom de champ d'index à mapper et le type de données du champ.

c. Choisissez Suivant.

9. Sur la page Réviser et créer, vérifiez que les informations saisies sont correctes, puis sélectionnez Ajouter une source de données. Vous pouvez également choisir de modifier vos informations depuis cette page. Votre source de données apparaîtra sur la page Sources de données une fois que la source de données aura été ajoutée avec succès.

API

Pour vous connecter Amazon Kendra à Amazon RDS (MySQL)

Vous devez spécifier les éléments suivants à l'aide de l'TemplateConfigurationAPI :

- Source de données —Spécifiez le type de source de données tel que `JDBC` lorsque vous utilisez le schéma TemplateConfigurationJSON. Spécifiez également la source de données `TEMPLATE` lorsque vous appelez l'CreateDataSourceAPI.
- Type de base de données —Vous devez spécifier le type de base de données sous `mySql` la forme.
- Requête SQL : spécifiez des instructions de requête SQL telles que les opérations SELECT et JOIN. La taille des requêtes SQL doit être inférieure à 32 Ko. Amazon Kendra analysera tout le contenu de la base de données correspondant à votre requête.
- Mode de synchronisation : spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation. Vous pouvez choisir entre :
  - `FORCED_FULL_CRAWL`pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
  - `FULL_CRAWL`pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
  - `CHANGE_LOG`pour indexer uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le

mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

- Nom de ressource Amazon (ARN) secret : indiquez le nom de ressource Amazon (ARN) d'un Secrets Manager secret contenant les informations d'authentification que vous avez créées dans votre Amazon RDS (MySQL) compte. Le secret est stocké dans une structure JSON avec les clés suivantes :

```
{
    "user name": "database user name",
    "password": "password"
}
```

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

- IAM role —Spécifiez le `RoleArn` moment où vous appelez `CreateDataSource` pour fournir à un IAM rôle les autorisations d'accéder à votre Secrets Manager secret et d'appeler le public requis APIs pour le Amazon RDS (MySQL) connecteur et Amazon Kendra. Pour plus d'informations, consultez la section IAM Rôles des sources de Amazon RDS (MySQL) données.

Vous pouvez également ajouter les fonctionnalités optionnelles suivantes :

- Virtual Private Cloud (VPC) **`VpcConfiguration`** —Spécifiez le moment de votre appel. `CreateDataSource` Pour de plus amples informations, veuillez consulter Configuration Amazon Kendra pour utiliser un Amazon VPC.
- Filtres d'inclusion et d'exclusion : vous pouvez spécifier si vous souhaitez inclure un contenu spécifique en utilisant l'utilisateur IDs, les groupes, la source URLs, les horodatages et les fuseaux horaires.
- Mappages de champs : choisissez de mapper les champs de votre source de Amazon RDS (MySQL) données à vos champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.

> **Note**
>
> Le champ du corps du document ou l'équivalent du corps du document pour vos documents est requis pour Amazon Kendra effectuer une recherche dans vos documents. Vous devez associer le nom du champ du corps du document dans votre source de données au nom du champ d'index`_document_body`. Tous les autres champs sont facultatifs.

- Filtrage du contexte utilisateur et contrôle d'accèsAmazon Kendra : analyse la liste de contrôle d'accès (ACL) de vos documents, si vous disposez d'une ACL pour vos documents. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.

Pour obtenir la liste des autres clés JSON importantes à configurer, consultezAmazon RDS Schéma de modèle (MySQL).

Remarques

- Les lignes de base de données supprimées ne seront pas suivies lors de la Amazon Kendra vérification du contenu mis à jour.
- La taille des noms et des valeurs des champs d'une ligne de votre base de données ne peut pas dépasser 400 Ko.
- Si la source de données de votre base de données contient une grande quantité de données et que vous ne souhaitez pas Amazon Kendra indexer tout le contenu de votre base de données après la première synchronisation, vous pouvez choisir de ne synchroniser que les documents nouveaux, modifiés ou supprimés.
- Il est recommandé de fournir des informations d'identification Amazon Kendra de base de données en lecture seule.
- Il est recommandé d'éviter d'ajouter des tableaux contenant des données sensibles ou des informations personnelles identifiables (PII).

Amazon RDS (Oracle)

> **Note**
>
> Amazon RDS (Oracle)Le connecteur reste entièrement pris en charge pour les clients existants jusqu'au 31 mai 2026. Bien que ce connecteur ne soit plus disponible pour les nouveaux utilisateurs, les utilisateurs actuels peuvent continuer à l'utiliser sans interruption. Nous faisons constamment évoluer notre portefeuille de connecteurs afin de proposer des solutions plus évolutives et personnalisables. Pour les intégrations futures, nous vous recommandons d'explorer le framework de connecteurs personnalisés Amazon Kendra [1], conçu pour prendre en charge un plus large éventail de cas d'utilisation en entreprise avec une flexibilité accrue.

Amazon RDS (Amazon Relational Database Service) est un service Web qui facilite la configuration, l'exploitation et le dimensionnement d'une base de données relationnelle dans AWS le cloud. Si vous êtes un Amazon RDS (Oracle) utilisateur, vous pouvez l'utiliser Amazon Kendra pour indexer votre source de Amazon RDS (Oracle) données. Le connecteur de source de Amazon Kendra Amazon RDS (Oracle) données prend en charge Amazon RDS Oracle Database 21c, Oracle Database 19c, Oracle Database 12c.

Vous pouvez vous connecter Amazon Kendra à votre source de Amazon RDS (Oracle) données à l'aide de la Amazon Kendra console et de l'TemplateConfigurationAPI.

Pour résoudre les problèmes liés à votre connecteur de source de Amazon Kendra Amazon RDS (Oracle) données, consultezDépannage des sources de données.

Rubriques

- Fonctionnalités prises en charge
- Prérequis
- Instructions de connexion
- Remarques

Fonctionnalités prises en charge

- Mappages de champs
- Filtrage du contexte utilisateur

- Filtres d'inclusion/exclusion
- Synchronisation complète et incrémentielle du contenu
- Cloud privé virtuel (VPC)

Prérequis

Avant de pouvoir Amazon Kendra indexer votre source de Amazon RDS (Oracle) données, apportez ces modifications à vos AWS comptes Amazon RDS (Oracle) and.

DansAmazon RDS (Oracle), assurez-vous d'avoir :

- Notez le nom d'utilisateur et le mot de passe de votre base de données.

> **Important**
>
> Il est recommandé de fournir des informations d'identification Amazon Kendra de base de données en lecture seule.

- Vous avez copié l'URL, le port et l'instance de votre hôte de base de données.
- Il est vérifié que chaque document est unique dans Amazon RDS (Oracle) et entre les autres sources de données que vous prévoyez d'utiliser pour le même index. Chaque source de données que vous souhaitez utiliser pour un index ne doit pas contenir le même document dans toutes les sources de données. IDs Les documents sont globaux par rapport à un index et doivent être uniques par index.

Dans votre Compte AWS, assurez-vous d'avoir :

- Création d'un Amazon Kendra index et, si vous utilisez l'API, notez l'ID de l'index.
- Vous avez créé un IAM rôle pour votre source de données et, si vous utilisez l'API, notez l'ARN du IAM rôle.

> **Note**
>
> Si vous modifiez votre type d'authentification et vos informations d'identification, vous devez mettre à jour votre IAM rôle pour accéder au bon identifiant AWS Secrets Manager secret.

- Stockez vos informations d'Amazon RDS (Oracle)authentification dans un AWS Secrets Manager secret et, si vous utilisez l'API, notez l'ARN du secret.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

Si vous n'avez pas de IAM rôle ou de secret existant, vous pouvez utiliser la console pour créer un nouveau IAM rôle et un nouveau Secrets Manager secret lorsque vous connectez votre source de Amazon RDS (Oracle) données à Amazon Kendra. Si vous utilisez l'API, vous devez fournir l'ARN d'un IAM rôle et d'un Secrets Manager secret existants, ainsi qu'un identifiant d'index.

Instructions de connexion

Pour vous connecter Amazon Kendra à votre source de Amazon RDS (Oracle) données, vous devez fournir les détails de vos Amazon RDS (Oracle) informations d'identification afin de Amazon Kendra pouvoir accéder à vos données. Si vous n'avez pas encore configuré Amazon RDS (Oracle) pour Amazon Kendra voirPrérequis.

Console

Pour vous connecter Amazon Kendra à Amazon RDS (Oracle)

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la.
2. Dans le volet de navigation de gauche, choisissez Index, puis choisissez l'index que vous souhaitez utiliser dans la liste des index.

> **Note**
>
> Vous pouvez choisir de configurer ou de modifier vos paramètres de contrôle d'accès utilisateur dans les paramètres de l'index.

3. Sur la page de démarrage, choisissez Ajouter une source de données.

4. Sur la page Ajouter une source de données, choisissez Amazon RDS (Oracle)connecteur, puis sélectionnez Ajouter un connecteur. Si vous utilisez la version 2 (le cas échéant), choisissez le Amazon RDS (Oracle)connecteur avec la balise « V2.0 ».
5. Sur la page Spécifier les détails de la source de données, entrez les informations suivantes :

   a. Dans Nom et description, pour Nom de la source de données : entrez le nom de votre source de données. Vous pouvez inclure des tirets, mais pas des espaces.

   b. (Facultatif) Description : entrez une description facultative pour votre source de données.

   c. Dans la langue par défaut : choisissez une langue pour filtrer vos documents pour l'index. Sauf indication contraire, la langue par défaut est l'anglais. La langue spécifiée dans les métadonnées du document remplace la langue sélectionnée.

   d. Dans Balises, pour Ajouter une nouvelle balise : incluez des balises facultatives pour rechercher et filtrer vos ressources ou suivre vos AWS coûts.

   e. Choisissez Suivant.
6. Sur la page Définir l'accès et la sécurité, entrez les informations suivantes :

   a. Dans Source, entrez les informations suivantes :

   b. Hôte : entrez le nom d'hôte de la base de données.

   c. Port — Entrez le port de base de données.

   d. Instance — Entrez l'instance de base de données.

   e. Activer l'emplacement du certificat SSL : choisissez d'entrer le Amazon S3 chemin d'accès à votre fichier de certificat SSL.

   f. Dans Authentification, entrez les informations suivantes :

      - AWS Secrets Manager secret —Choisissez un secret existant ou créez-en un nouveau Secrets Manager pour stocker vos informations Amazon RDS (Oracle) d'authentification. Si vous choisissez de créer un nouveau secret, une fenêtre AWS Secrets Manager secrète s'ouvre.

        A. Entrez les informations suivantes dans la fenêtre Créer un AWS Secrets Manager secret :

           I. Nom secret : le nom de votre secret. Le préfixe « AmazonKendra - Amazon RDS (Oracle) - » est automatiquement ajouté à votre nom secret.

II. Pour le nom d'utilisateur et le mot de passe de la base de données, entrez les valeurs d'identification d'authentification que vous avez copiées depuis votre base de données.

B. Choisissez Enregistrer.

g. Virtual Private Cloud (VPC) —Vous pouvez choisir d'utiliser un VPC. Dans ce cas, vous devez ajouter des sous-réseaux et des groupes de sécurité VPC.

h. IAM rôle —Choisissez un IAM rôle existant ou créez-en un nouveau IAM pour accéder aux informations d'identification de votre référentiel et indexer le contenu.

> **Note**
>
> IAM les rôles utilisés pour les index ne peuvent pas être utilisés pour les sources de données. Si vous ne savez pas si un rôle existant est utilisé pour un index ou une FAQ, choisissez Créer un nouveau rôle pour éviter les erreurs.

i. Choisissez Suivant.

7. Sur la page Configurer les paramètres de synchronisation, entrez les informations suivantes :

a. Dans le champ d'application de la synchronisation, choisissez l'une des options suivantes :

- Requête SQL —Entrez des instructions de requête SQL telles que les opérations SELECT et JOIN. La taille des requêtes SQL doit être inférieure à 32 Ko. Amazon Kendra analysera tout le contenu de la base de données correspondant à votre requête.
- Colonne clé primaire : indiquez la clé primaire pour la table de base de données. Cela identifie une table dans votre base de données.
- Colonne de titre : indiquez le nom de la colonne de titre du document dans votre table de base de données.
- Colonne du corps : indiquez le nom de la colonne du corps du document dans votre table de base de données.

b. Dans Configuration supplémentaire — facultatif, choisissez l'une des options suivantes pour synchroniser un contenu spécifique au lieu de synchroniser tous les fichiers :

- Colonnes détectant les modifications : entrez le nom des colonnes qui Amazon Kendra seront utilisées pour détecter les modifications de contenu. Amazon Kendra réindexera le contenu en cas de modification de l'une de ces colonnes.
- IDsColonne des utilisateurs : entrez le nom de la colonne contenant l'utilisateur devant être autorisé IDs à accéder au contenu.
- Colonne Groupes : entrez le nom de la colonne contenant les groupes autorisés à accéder au contenu.
- URLsColonne source —Entrez le nom de la colonne contenant la source URLs à indexer.
- Colonne horodatage : entrez le nom de la colonne contenant les horodatages. Amazon Kendra utilise les informations d'horodatage pour détecter les modifications apportées à votre contenu et synchroniser uniquement le contenu modifié.
- Colonne fuseaux horaires : entrez le nom de la colonne contenant les fuseaux horaires du contenu à analyser.
- Format d'horodatage : entrez le nom de la colonne contenant les formats d'horodatage à utiliser pour détecter les modifications de contenu et resynchroniser votre contenu.

c. Mode de synchronisation : choisissez la manière dont vous souhaitez mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation.

- Synchronisation complète : Indexez tout le contenu fraîchement, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
- Nouvelle synchronisation modifiée : Indexez uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
- Synchronisation nouvelle, modifiée ou supprimée : Indexez uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre

source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

d. Dans Synchroniser le calendrier d'exécution, pour Fréquence : à quelle fréquence Amazon Kendra sera synchronisée avec votre source de données.

e. Choisissez Suivant.

8. Sur la page Définir les mappages de champs, entrez les informations suivantes :

a. Sélectionnez parmi les champs de source de données par défaut générés (Document IDs, Titres des documents et Source URLs) que vous souhaitez Amazon Kendra mapper à l'index.

b. Ajouter un champ —Pour ajouter des champs de source de données personnalisés afin de créer un nom de champ d'index à mapper et le type de données du champ.

c. Choisissez Suivant.

9. Sur la page Réviser et créer, vérifiez que les informations saisies sont correctes, puis sélectionnez Ajouter une source de données. Vous pouvez également choisir de modifier vos informations depuis cette page. Votre source de données apparaîtra sur la page Sources de données une fois que la source de données aura été ajoutée avec succès.

API

Pour vous connecter Amazon Kendra à Amazon RDS (Oracle)

Vous devez spécifier les éléments suivants à l'aide de l'TemplateConfigurationAPI :

- Source de données —Spécifiez le type de source de données tel que `JDBC` lorsque vous utilisez le schéma TemplateConfigurationJSON. Spécifiez également la source de données `TEMPLATE` lorsque vous appelez l'CreateDataSourceAPI.
- Type de base de données —Vous devez spécifier le type de base de données sous `oracle` la forme.
- Requête SQL : spécifiez des instructions de requête SQL telles que les opérations SELECT et JOIN. La taille des requêtes SQL doit être inférieure à 32 Ko. Amazon Kendra analysera tout le contenu de la base de données correspondant à votre requête.
- Mode de synchronisation : spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la

synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation. Vous pouvez choisir entre :

- `FORCED_FULL_CRAWL`pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
- `FULL_CRAWL`pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
- `CHANGE_LOG`pour indexer uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

- Nom de ressource Amazon (ARN) secret : indiquez le nom de ressource Amazon (ARN) d'un Secrets Manager secret contenant les informations d'authentification que vous avez créées dans votre Amazon RDS (Oracle) compte. Le secret est stocké dans une structure JSON avec les clés suivantes :

```
{
    "user name": "database user name",
    "password": "password"
}
```

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

- IAM role —Spécifiez à quel `RoleArn` moment vous appelez `CreateDataSource` pour fournir à un IAM rôle les autorisations d'accéder à votre Secrets Manager secret et d'appeler le public requis APIs pour le Amazon RDS (Oracle) connecteur et Amazon Kendra. Pour plus d'informations, consultez la section IAM Rôles des sources de Amazon RDS (Oracle) données.

Vous pouvez également ajouter les fonctionnalités optionnelles suivantes :

- Virtual Private Cloud (VPC) **`VpcConfiguration`** —Spécifiez le moment de votre appel. `CreateDataSource` Pour de plus amples informations, veuillez consulter Configuration Amazon Kendra pour utiliser un Amazon VPC.
- Filtres d'inclusion et d'exclusion : vous pouvez spécifier si vous souhaitez inclure un contenu spécifique en utilisant l'utilisateur IDs, les groupes, la source URLs, les horodatages et les fuseaux horaires.
- Filtrage du contexte utilisateur et contrôle d'accèsAmazon Kendra : analyse la liste de contrôle d'accès (ACL) de vos documents, si vous disposez d'une ACL pour vos documents. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.
- Mappages de champs : choisissez de mapper les champs de votre source de Amazon RDS (Oracle) données à vos champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.

> **Note**
>
> Le champ du corps du document ou l'équivalent du corps du document pour vos documents est requis pour Amazon Kendra effectuer une recherche dans vos documents. Vous devez associer le nom du champ du corps du document dans votre source de données au nom du champ d'index`_document_body`. Tous les autres champs sont facultatifs.

Pour obtenir la liste des autres clés JSON importantes à configurer, consultezAmazon RDS schéma de modèle (Oracle).

Remarques

- Les lignes de base de données supprimées ne seront pas suivies lors de la Amazon Kendra vérification du contenu mis à jour.
- La taille des noms et des valeurs des champs d'une ligne de votre base de données ne peut pas dépasser 400 Ko.
- Si la source de données de votre base de données contient une grande quantité de données et que vous ne Amazon Kendra souhaitez pas indexer tout le contenu de votre base de données

après la première synchronisation, vous pouvez choisir de synchroniser uniquement les documents nouveaux, modifiés ou supprimés.

- Il est recommandé de fournir des informations d'identification Amazon Kendra de base de données en lecture seule.
- Il est recommandé d'éviter d'ajouter des tableaux contenant des données sensibles ou des informations personnelles identifiables (PII).

Amazon RDS (PostgreSQL)

> **Note**
>
> Amazon RDS (PostgreSQL)Le connecteur reste entièrement pris en charge pour les clients existants jusqu'au 31 mai 2026. Bien que ce connecteur ne soit plus disponible pour les nouveaux utilisateurs, les utilisateurs actuels peuvent continuer à l'utiliser sans interruption. Nous faisons constamment évoluer notre portefeuille de connecteurs afin de proposer des solutions plus évolutives et personnalisables. Pour les intégrations futures, nous vous recommandons d'explorer le framework de connecteurs personnalisés Amazon Kendra [1], conçu pour prendre en charge un plus large éventail de cas d'utilisation en entreprise avec une flexibilité accrue.

Amazon RDS est un service Web qui facilite la configuration, l'exploitation et le dimensionnement d'une base de données relationnelle dans le AWS cloud. Si vous êtes un Amazon RDS utilisateur, vous pouvez l'utiliser Amazon Kendra pour indexer votre source de Amazon RDS (PostgreSQL) données. Le connecteur de source de Amazon Kendra Amazon RDS (PostgreSQL) données prend en charge PostgreSQL 9.6.

Vous pouvez vous connecter Amazon Kendra à votre source de Amazon RDS (PostgreSQL) données à l'aide de la Amazon Kendra console et de l'TemplateConfigurationAPI.

Pour résoudre les problèmes liés à votre connecteur de source de Amazon Kendra Amazon RDS (PostgreSQL) données, consultezDépannage des sources de données.

Rubriques

- Fonctionnalités prises en charge
- Prérequis

- Instructions de connexion
- Remarques

Fonctionnalités prises en charge

- Mappages de champs
- Filtrage du contexte utilisateur
- Filtres d'inclusion/exclusion
- Synchronisation complète et incrémentielle du contenu
- Cloud privé virtuel (VPC)

Prérequis

Avant de pouvoir Amazon Kendra indexer votre source de Amazon RDS (PostgreSQL) données, apportez ces modifications à vos AWS comptes Amazon RDS (PostgreSQL) and.

DansAmazon RDS (PostgreSQL), assurez-vous d'avoir :

- Notez le nom d'utilisateur et le mot de passe de votre base de données.

  > **Important**
  >
  > Il est recommandé de fournir des informations d'identification Amazon Kendra de base de données en lecture seule.

- Vous avez copié l'URL, le port et l'instance de votre hôte de base de données. Vous pouvez trouver ces informations sur la Amazon RDS console.
- Il est vérifié que chaque document est unique dans Amazon RDS (PostgreSQL) et entre les autres sources de données que vous prévoyez d'utiliser pour le même index. Chaque source de données que vous souhaitez utiliser pour un index ne doit pas contenir le même document dans toutes les sources de données. IDs Les documents sont globaux par rapport à un index et doivent être uniques par index.

Dans votre Compte AWS, assurez-vous d'avoir :

- Création d'un Amazon Kendra index et, si vous utilisez l'API, notez l'ID de l'index.

- Vous avez créé un IAM rôle pour votre source de données et, si vous utilisez l'API, notez l'ARN du IAM rôle.

> **Note**
>
> Si vous modifiez votre type d'authentification et vos informations d'identification, vous devez mettre à jour votre IAM rôle pour accéder au bon identifiant AWS Secrets Manager secret.

- Stockez vos informations d'Amazon RDS (PostgreSQL)authentification dans un AWS Secrets Manager secret et, si vous utilisez l'API, notez l'ARN du secret.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

Si vous n'avez pas de IAM rôle ou de secret existant, vous pouvez utiliser la console pour créer un nouveau IAM rôle et un nouveau Secrets Manager secret lorsque vous connectez votre source de Amazon RDS (PostgreSQL) données à Amazon Kendra. Si vous utilisez l'API, vous devez fournir l'ARN d'un IAM rôle et d'un Secrets Manager secret existants, ainsi qu'un identifiant d'index.

Instructions de connexion

Pour vous connecter Amazon Kendra à votre source de Amazon RDS (PostgreSQL) données, vous devez fournir les détails de vos Amazon RDS (PostgreSQL) informations d'identification afin de Amazon Kendra pouvoir accéder à vos données. Si vous n'avez pas encore configuré Amazon RDS (PostgreSQL) pour Amazon Kendra voirPrérequis.

Console

Pour vous connecter Amazon Kendra à Amazon RDS (PostgreSQL)

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la.

2. Dans le volet de navigation de gauche, choisissez Index, puis choisissez l'index que vous souhaitez utiliser dans la liste des index.

> **Note**
>
> Vous pouvez choisir de configurer ou de modifier vos paramètres de contrôle d'accès utilisateur dans les paramètres de l'index.

3. Sur la page de démarrage, choisissez Ajouter une source de données.
4. Sur la page Ajouter une source de données, choisissez Amazon RDS (PostgreSQL)connecteur, puis sélectionnez Ajouter un connecteur. Si vous utilisez la version 2 (le cas échéant), choisissez le Amazon RDS (PostgreSQL)connecteur avec la balise « V2.0 ».
5. Sur la page Spécifier les détails de la source de données, entrez les informations suivantes :
   a. Dans Nom et description, pour Nom de la source de données : entrez le nom de votre source de données. Vous pouvez inclure des traits d'union, mais pas des espaces.
   b. (Facultatif) Description : entrez une description facultative pour votre source de données.
   c. Dans la langue par défaut : choisissez une langue pour filtrer vos documents pour l'index. Sauf indication contraire, la langue par défaut est l'anglais. La langue spécifiée dans les métadonnées du document remplace la langue sélectionnée.
   d. Dans Balises, pour Ajouter une nouvelle balise : incluez des balises facultatives pour rechercher et filtrer vos ressources ou suivre vos AWS coûts.
   e. Choisissez Suivant.
6. Sur la page Définir l'accès et la sécurité, entrez les informations suivantes :
   a. Dans Source, entrez les informations suivantes :
   b. Hôte : entrez l'URL de l'hôte de la base de données, par exemple :`http://`*`instance URL`*`.`*`region`*`.rds.amazonaws.com`.
   c. Port — Entrez le port de base de données, par exemple,`5432`.
   d. Instance — Entrez l'instance de base de données, par exemple`postgres`.
   e. Activer l'emplacement du certificat SSL : choisissez d'entrer le Amazon S3 chemin d'accès à votre fichier de certificat SSL.
   f. Dans Authentification, entrez les informations suivantes :

- AWS Secrets Manager secret —Choisissez un secret existant ou créez-en un nouveau Secrets Manager pour stocker vos informations Amazon RDS (PostgreSQL) d'authentification. Si vous choisissez de créer un nouveau secret, une fenêtre AWS Secrets Manager secrète s'ouvre.

  A. Entrez les informations suivantes dans la fenêtre Créer un AWS Secrets Manager secret :

     I. Nom secret : le nom de votre secret. Le préfixe « AmazonKendra - Amazon RDS (PostgreSQL) - » est automatiquement ajouté à votre nom secret.

     II. Pour le nom d'utilisateur et le mot de passe de la base de données, entrez les valeurs d'identification d'authentification que vous avez copiées depuis votre base de données.

  B. Choisissez Enregistrer.

g. Virtual Private Cloud (VPC) —Vous pouvez choisir d'utiliser un VPC. Dans ce cas, vous devez ajouter des sous-réseaux et des groupes de sécurité VPC.

h. IAM rôle —Choisissez un IAM rôle existant ou créez-en un nouveau IAM pour accéder aux informations d'identification de votre référentiel et indexer le contenu.

> **Note**
>
> IAM les rôles utilisés pour les index ne peuvent pas être utilisés pour les sources de données. Si vous ne savez pas si un rôle existant est utilisé pour un index ou une FAQ, choisissez Créer un nouveau rôle pour éviter les erreurs.

i. Choisissez Suivant.

7. Sur la page Configurer les paramètres de synchronisation, entrez les informations suivantes :

   a. Dans le champ d'application de la synchronisation, choisissez l'une des options suivantes :

   - Requête SQL —Entrez des instructions de requête SQL telles que les opérations SELECT et JOIN. Les requêtes SQL doivent être inférieures à 32 Ko Les requêtes SQL doivent être inférieures à 32 Ko et ne pas contenir de points-virgules (;). Amazon Kendra analysera tout le contenu de la base de données correspondant à votre requête.

- Colonne clé primaire : indiquez la clé primaire pour la table de base de données. Cela identifie une table au sein de votre base de données.
- Colonne de titre : indiquez le nom de la colonne de titre du document dans votre table de base de données.
- Colonne du corps : indiquez le nom de la colonne du corps du document dans votre table de base de données.

b. Dans Configuration supplémentaire — facultatif, choisissez l'une des options suivantes pour synchroniser un contenu spécifique au lieu de synchroniser tous les fichiers :

- Colonnes détectant les modifications : entrez le nom des colonnes qui Amazon Kendra seront utilisées pour détecter les modifications de contenu. Amazon Kendra réindexera le contenu en cas de modification de l'une de ces colonnes.
- IDsColonne des utilisateurs : entrez le nom de la colonne contenant l'utilisateur devant être autorisé IDs à accéder au contenu.
- Colonne Groupes : entrez le nom de la colonne qui contient les groupes autorisés à accéder au contenu.
- URLsColonne source —Entrez le nom de la colonne contenant la source URLs à indexer.
- Colonne horodatage : entrez le nom de la colonne contenant les horodatages. Amazon Kendra utilise les informations d'horodatage pour détecter les modifications apportées à votre contenu et synchroniser uniquement le contenu modifié.
- Colonne fuseaux horaires : entrez le nom de la colonne contenant les fuseaux horaires du contenu à analyser.
- Format d'horodatage : entrez le nom de la colonne contenant les formats d'horodatage à utiliser pour détecter les modifications de contenu et resynchroniser votre contenu.

c. Mode de synchronisation : choisissez la manière dont vous souhaitez mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation.

- Synchronisation complète : Indexez tout le contenu fraîchement, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.

- Nouvelle synchronisation modifiée : Indexez uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
- Synchronisation nouvelle, modifiée ou supprimée : Indexez uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

d. Dans Synchroniser le calendrier d'exécution, pour Fréquence : à quelle fréquence Amazon Kendra se synchronisera avec votre source de données.

e. Choisissez Suivant.

8. Sur la page Définir les mappages de champs, entrez les informations suivantes :

   a. Sélectionnez parmi les champs de source de données par défaut générés (Document IDs, Titres des documents et Source URLs) que vous souhaitez Amazon Kendra mapper à l'index.

   b. Ajouter un champ —Pour ajouter des champs de source de données personnalisés afin de créer un nom de champ d'index à mapper et le type de données du champ.

   c. Choisissez Suivant.

9. Sur la page Réviser et créer, vérifiez que les informations saisies sont correctes, puis sélectionnez Ajouter une source de données. Vous pouvez également choisir de modifier vos informations à partir de cette page. Votre source de données apparaîtra sur la page Sources de données une fois que la source de données aura été ajoutée avec succès.

API

Pour vous connecter Amazon Kendra à Amazon RDS (PostgreSQL)

Vous devez spécifier les éléments suivants à l'aide de l'TemplateConfigurationAPI :

- Source de données —Spécifiez le type de source de données tel que `JDBC` lorsque vous utilisez le schéma TemplateConfigurationJSON. Spécifiez également la source de données `TEMPLATE` lorsque vous appelez l'CreateDataSourceAPI.

- Type de base de données —Vous devez spécifier le type de base de données sous `postgresql` la forme.
- Requête SQL : spécifiez des instructions de requête SQL telles que les opérations SELECT et JOIN. La taille des requêtes SQL doit être inférieure à 32 Ko. Amazon Kendra analysera tout le contenu de la base de données correspondant à votre requête.
- Mode de synchronisation : spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation. Vous pouvez choisir entre :
  - `FORCED_FULL_CRAWL`pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
  - `FULL_CRAWL`pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
  - `CHANGE_LOG`pour indexer uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
- Nom de ressource Amazon (ARN) secret : indiquez le nom de ressource Amazon (ARN) d'un Secrets Manager secret contenant les informations d'authentification que vous avez créées dans votre Amazon RDS (PostgreSQL) compte. Le secret est stocké dans une structure JSON avec les clés suivantes :

```
{
    "user name": "database user name",
    "password": "password"
}
```

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les

informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

- IAM role —Spécifiez à quel `RoleArn` moment vous appelez `CreateDataSource` pour fournir à un IAM rôle les autorisations d'accéder à votre Secrets Manager secret et d'appeler le public requis APIs pour le Amazon RDS (PostgreSQL) connecteur et Amazon Kendra. Pour plus d'informations, consultez la section IAM Rôles des sources de Amazon RDS (PostgreSQL) données.

Vous pouvez également ajouter les fonctionnalités optionnelles suivantes :

- Virtual Private Cloud (VPC) **`VpcConfiguration`** —Spécifiez le moment de votre appel. `CreateDataSource` Pour de plus amples informations, veuillez consulter Configuration Amazon Kendra pour utiliser un Amazon VPC.
- Filtres d'inclusion et d'exclusion : vous pouvez spécifier si vous souhaitez inclure un contenu spécifique en utilisant l'utilisateur IDs, les groupes, la source URLs, les horodatages et les fuseaux horaires.
- Filtrage du contexte utilisateur et contrôle d'accèsAmazon Kendra : analyse la liste de contrôle d'accès (ACL) de vos documents, si vous disposez d'une ACL pour vos documents. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.
- Mappages de champs : choisissez de mapper les champs de votre source de Amazon RDS (PostgreSQL) données à vos champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.

> **Note**
>
> Le champ du corps du document ou l'équivalent du corps du document pour vos documents est requis pour Amazon Kendra effectuer une recherche dans vos documents. Vous devez associer le nom du champ du corps du document dans votre source de données au nom du champ d'index`_document_body`. Tous les autres champs sont facultatifs.

Pour obtenir la liste des autres clés JSON importantes à configurer, consultezAmazon RDS Schéma de modèle (PostgreSQL).

Remarques

- Les lignes de base de données supprimées ne seront pas suivies lors de la Amazon Kendra vérification du contenu mis à jour.
- La taille des noms et des valeurs des champs d'une ligne de votre base de données ne peut pas dépasser 400 Ko.
- Si la source de données de votre base de données contient une grande quantité de données et que vous ne souhaitez pas Amazon Kendra indexer tout le contenu de votre base de données après la première synchronisation, vous pouvez choisir de ne synchroniser que les documents nouveaux, modifiés ou supprimés.
- Il est recommandé de fournir des informations d'identification Amazon Kendra de base de données en lecture seule.
- Il est recommandé d'éviter d'ajouter des tableaux contenant des données sensibles ou des informations personnelles identifiables (PII).

Amazon S3

Amazon S3 est un service de stockage d'objets qui stocke les données sous forme d'objets dans des compartiments. Vous pouvez l'utiliser Amazon Kendra pour indexer le référentiel de documents de votre Amazon S3 bucket.

> **Warning**
>
> Amazon Kendra n'utilise pas de politique de compartiment qui autorise un Amazon Kendra mandant à interagir avec un compartiment S3. Au lieu de cela, il utilise IAM des rôles. Assurez-vous qu'il Amazon Kendra n'est pas inclus en tant que membre de confiance dans votre politique de compartiment afin d'éviter tout problème de sécurité des données lié à l'octroi accidentel d'autorisations à des principaux arbitraires. Vous pouvez toutefois ajouter une politique de compartiment pour utiliser un Amazon S3 compartiment sur différents comptes. Pour plus d'informations, consultez la section Politiques à Amazon S3 utiliser entre les comptes (dans l'onglet IAM Rôles S3, sous IAM Rôles pour les sources de données).

Pour plus d'informations sur IAM les rôles pour les sources de données S3, consultez la section IAM rôles.

> **Note**
>
> Amazon Kendra prend désormais en charge un Amazon S3 connecteur amélioré. La console a été automatiquement mise à niveau pour vous. Tous les nouveaux connecteurs que vous créez dans la console utiliseront l'architecture mise à niveau. Si vous utilisez l'API, vous devez désormais utiliser l'TemplateConfigurationobjet au lieu de l'`S3DataSourceConfiguration`objet pour configurer votre connecteur.
> Les connecteurs configurés à l'aide de l'ancienne console et de l'ancienne architecture d'API continueront de fonctionner tels qu'ils ont été configurés. Toutefois, vous ne pourrez ni les modifier ni les mettre à jour. Si vous souhaitez modifier ou mettre à jour la configuration de votre connecteur, vous devez créer un nouveau connecteur.
> Nous vous recommandons de migrer le flux de travail de votre connecteur vers la version mise à niveau. Support pour les connecteurs configurés à l'aide de l'ancienne architecture devrait prendre fin en juin 2024.

Vous pouvez vous connecter à votre source de Amazon S3 données à l'aide de la Amazon Kendra console ou de l'TemplateConfigurationAPI.

> **Note**
>
> Pour générer un rapport d'état de synchronisation pour votre source de Amazon S3 données, consultez la section Résolution des problèmes liés aux sources de données.

Pour résoudre les problèmes liés à votre connecteur de source de données Amazon Kendra S3, consultezDépannage des sources de données.

Rubriques

- Fonctionnalités prises en charge
- Prérequis
- Instructions de connexion
- Création d'une source Amazon S3 de données

- Amazon S3 métadonnées du document
- Contrôle d'accès pour les sources Amazon S3 de données
- Utilisation Amazon VPC avec une source Amazon S3 de données

Fonctionnalités prises en charge

- Mappages de champs
- Contrôle d'accès utilisateur
- Filtres d'inclusion/exclusion
- Synchronisation complète et incrémentielle du contenu
- Cloud privé virtuel (VPC)

Prérequis

Avant de pouvoir l'utiliser Amazon Kendra pour indexer votre source de données S3, apportez ces modifications à votre S3 et à vos AWS comptes.

Dans S3, assurez-vous d'avoir :

- Vous avez copié le nom de votre Amazon S3 compartiment.

  > **Note**
  >
  > Votre compartiment doit se trouver dans la même région que votre Amazon Kendra index et celui-ci doit être autorisé à accéder au compartiment contenant vos documents.

- Il est vérifié que chaque document est unique dans S3 et dans les autres sources de données que vous prévoyez d'utiliser pour le même index. Chaque source de données que vous souhaitez utiliser pour un index ne doit pas contenir le même document dans toutes les sources de données. IDs Les documents sont globaux par rapport à un index et doivent être uniques par index.

Dans votre AWS compte, assurez-vous d'avoir :

- Création d'un Amazon Kendra index et, si vous utilisez l'API, notez l'ID de l'index.
- Vous avez créé un IAM rôle pour votre source de données et, si vous utilisez l'API, notez l'ARN du IAM rôle.

Si vous n'avez pas de IAM rôle existant, vous pouvez utiliser la console pour créer un nouveau IAM rôle lorsque vous connectez votre source de données S3 à Amazon Kendra. Si vous utilisez l'API, vous devez fournir l'ARN d'un IAM rôle existant et un identifiant d'index.

Instructions de connexion

Pour vous connecter Amazon Kendra à votre source de données S3, vous devez fournir les informations nécessaires sur votre source de données S3 afin de Amazon Kendra pouvoir accéder à vos données. Si vous n'avez pas encore configuré S3 pour Amazon Kendra, consultezPrérequis.

Console

Pour vous connecter Amazon Kendra à Amazon S3

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la.
2. Dans le volet de navigation de gauche, choisissez Index, puis choisissez l'index que vous souhaitez utiliser dans la liste des index.

   > **Note**
   >
   > Vous pouvez choisir de configurer ou de modifier vos paramètres de contrôle d'accès utilisateur dans les paramètres de l'index.

3. Sur la page de démarrage, choisissez Ajouter une source de données.
4. Sur la page Ajouter une source de données, choisissez le connecteur S3, puis choisissez Ajouter un connecteur. Si vous utilisez la version 2 (le cas échéant), choisissez le connecteur S3 avec la balise « V2.0 ».
5. Sur la page Spécifier les détails de la source de données, entrez les informations suivantes :

   a. Dans Nom et description, pour Nom de la source de données : entrez le nom de votre source de données. Vous pouvez inclure des traits d'union, mais pas des espaces.
   b. (Facultatif) Description : entrez une description facultative pour votre source de données.
   c. Dans la langue par défaut : choisissez une langue pour filtrer vos documents pour l'index. Sauf indication contraire, la langue par défaut est l'anglais. La langue spécifiée dans les métadonnées du document remplace la langue sélectionnée.
   d. Dans Balises, pour Ajouter une nouvelle balise : incluez des balises facultatives pour rechercher et filtrer vos ressources ou suivre vos AWS coûts.
   e. Choisissez Suivant.

6. Sur la page Définir l'accès et la sécurité, entrez les informations facultatives suivantes :

   a. IAM rôle —Choisissez un IAM rôle existant ou créez-en un nouveau IAM pour accéder aux informations d'identification de votre référentiel et indexer le contenu.

   > **Note**
   >
   > IAM les rôles utilisés pour les index ne peuvent pas être utilisés pour les sources de données. Si vous ne savez pas si un rôle existant est utilisé pour un index ou une FAQ, choisissez Créer un nouveau rôle pour éviter les erreurs.

   b. Virtual Private Cloud (VPC) —Vous pouvez choisir d'utiliser un VPC. Dans ce cas, vous devez ajouter des sous-réseaux et des groupes de sécurité VPC.

   c. Choisissez Suivant.

7. Sur la page Configurer les paramètres de synchronisation, entrez les informations suivantes :

   a. Pour l'emplacement de la source de données : spécifiez le chemin d'accès au Amazon S3 compartiment dans lequel vos données sont stockées. Sélectionnez Parcourir S3 pour choisir votre compartiment S3.

   b. Pour la taille de fichier maximale : spécifiez une limite en Mo pour analyser uniquement les fichiers inférieurs à cette limite. La taille de fichier maximale Amazon Kendra autorisée est de 50 Mo.

   c. Pour les fichiers de métadonnées (facultatif), préfixez l'emplacement du dossier : spécifiez le chemin d'accès au dossier dans lequel vos métadonnées fields/attributes et celles des autres documents sont stockées. Sélectionnez Parcourir S3 pour localiser votre dossier de métadonnées.

   d. Pour l'emplacement du fichier de configuration de la liste de contrôle d'accès (facultatif) : spécifiez le chemin d'accès au fichier contenant une structure JSON de vos utilisateurs et de leur accès aux documents. Sélectionnez Parcourir S3 pour localiser votre fichier ACL.

   e. (Facultatif) Sélectionnez la clé de déchiffrement : sélectionnez cette option pour utiliser une clé de déchiffrement. Vous pouvez choisir d'utiliser une AWS KMS clé existante.

   f. Pour une configuration supplémentaire (facultative) : ajoutez des modèles pour inclure ou exclure certains fichiers. Tous les chemins sont relatifs au compartiment S3 de l'emplacement de la source de données.

g. Mode de synchronisation : choisissez la manière dont vous souhaitez mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation.

   - Synchronisation complète : Indexez tout le contenu fraîchement, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
   - Synchronisation nouvelle, modifiée ou supprimée : Indexez uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

h. Dans Calendrier d'exécution de la synchronisation, pour Fréquence : choisissez la fréquence à laquelle vous souhaitez synchroniser le contenu de votre source de données et mettre à jour votre index.

i. Choisissez Suivant.

8. Sur la page Définir les mappages de champs, entrez les informations facultatives suivantes :

   a. Mappages de champs par défaut : sélectionnez parmi les champs de source de données par défaut Amazon Kendra générés que vous souhaitez mapper à votre index.

   b. Ajouter un champ : choisissez d'ajouter des champs de source de données personnalisés pour créer un nom de champ d'index à mapper et le type de données du champ.

   c. Choisissez Suivant.

9. Sur la page Réviser et créer, vérifiez que les informations saisies sont correctes, puis sélectionnez Ajouter une source de données. Vous pouvez également choisir de modifier vos informations depuis cette page. Votre source de données apparaîtra sur la page Sources de données une fois que la source de données aura été ajoutée avec succès.

API

Pour vous connecter Amazon Kendra à Amazon S3

Vous devez spécifier un JSON du schéma de source de données à l'aide de l'TemplateConfigurationAPI. Vous devez fournir les informations suivantes :

- Source de données —Spécifiez le type de source de données tel que `S3` lorsque vous utilisez le schéma TemplateConfigurationJSON. Spécifiez également la source de données `TEMPLATE` lorsque vous appelez l'CreateDataSourceAPI.
- BucketName: le nom du compartiment contenant les documents.
- Mode de synchronisation : spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation. Vous pouvez choisir entre :
  - `FORCED_FULL_CRAWL`pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
  - `FULL_CRAWL`pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
- IAM role —Spécifiez à quel `RoleArn` moment vous appelez `CreateDataSource` pour fournir à un IAM rôle les autorisations d'accéder à votre Secrets Manager secret et d'appeler le public requis APIs pour le connecteur S3 et Amazon Kendra. Pour plus d'informations, consultez la section IAM Rôles des sources de données S3.

Vous pouvez également ajouter les fonctionnalités optionnelles suivantes :

- Virtual Private Cloud (VPC) **`VpcConfiguration`** —Spécifiez le moment de votre appel. `CreateDataSource` Pour de plus amples informations, veuillez consulter Configuration Amazon Kendra pour utiliser un Amazon VPC.
- Filtres d'inclusion et d'exclusion : spécifiez s'il faut inclure ou exclure certains noms de fichiers, types de fichiers et chemins de fichiers. Vous utilisez des modèles globulaires (modèles qui peuvent transformer un motif générique en une liste de noms de chemins correspondant au modèle donné). Pour des exemples, consultez la section Utilisation des filtres d'exclusion et d'inclusion dans la référence des commandes de la AWS CLI.

- Configuration des métadonnées et du contrôle d'accès aux documents : ajoutez des métadonnées de document et des fichiers de contrôle d'accès contenant des informations telles que l'URI source, l'auteur du document ou les attributs/champs personnalisés du document, ainsi que vos utilisateurs et les documents auxquels ils peuvent accéder. Chaque fichier de métadonnées contient des métadonnées relatives à un seul document.
- Mappages de champs : choisissez de mapper les champs de votre source de données S3 à vos champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.

  > **Note**
  >
  > Le champ du corps du document ou l'équivalent du corps du document pour vos documents est requis pour Amazon Kendra effectuer une recherche dans vos documents. Vous devez associer le nom du champ du corps du document dans votre source de données au nom du champ d'index`_document_body`. Tous les autres champs sont facultatifs.

  Pour obtenir la liste des autres clés JSON importantes à configurer, consultez le schéma du S3 modèle.

En savoir plus

Pour en savoir plus sur l'intégration Amazon Kendra à votre source de données S3, consultez :

- Recherchez des réponses avec précision à l'aide du connecteur Amazon Kendra S3 avec support VPC

Création d'une source Amazon S3 de données

Les exemples suivants illustrent la création d'une source de Amazon S3 données. Les exemples supposent que vous avez déjà créé un index et un IAM rôle autorisés à lire les données de l'index. Pour plus d'informations sur le IAM rôle, consultez la section Rôles IAM d'accès. Pour plus d'informations sur la création d'un index, consultez la section Création d'un index.

CLI

```
aws kendra create-data-source \
```

```
--index-id index ID \
--name example-data-source \
--type S3 \
--configuration '{"S3Configuration":{"BucketName":"bucket name"}}'
--role-arn 'arn:aws:iam::account id:role:/role name
```

Python

L'extrait de code Python suivant crée une source de Amazon S3 données. Pour un exemple complet, voirDémarrer (AWS SDK pour Python (Boto3)).

```
print("Create an Amazon S3 data source.")

    # Provide a name for the data source
    name = "getting-started-data-source"
    # Provide an optional description for the data source
    description = "Getting started data source."
    # Provide the IAM role ARN required for data sources
    role_arn = "arn:aws:iam::${accountID}:role/${roleName}"
    # Provide the data soource connection information
    s3_bucket_name = "S3-bucket-name"
    type = "S3"
    # Configure the data source
    configuration = {"S3DataSourceConfiguration":
        {
            "BucketName": s3_bucket_name
        }
    }

    data_source_response = kendra.create_data_source(
        Configuration = configuration,
        Name = name,
        Description = description,
        RoleArn = role_arn,
        Type = type,
        IndexId = index_id
    )
```

La création de votre source de données peut prendre un certain temps. Vous pouvez suivre la progression à l'aide de l'DescribeDataSourceAPI. Lorsque le statut de la source de données est `ACTIVE` défini, la source de données est prête à être utilisée.

Les exemples suivants montrent comment obtenir le statut d'une source de données.

CLI

```
aws kendra describe-data-source \
 --index-id index ID \
 --id data source ID
```

Python

L'extrait de code Python suivant fournit des informations sur une source de données S3. Pour un exemple complet, voirDémarrer (AWS SDK pour Python (Boto3)).

```
print("Wait for Amazon Kendra to create the data source.")

    while True:
        data_source_description = kendra.describe_data_source(
            Id = "data-source-id",
            IndexId = "index-id"
        )
        status = data_source_description["Status"]
        print(" Creating data source. Status: "+status)
        time.sleep(60)
        if status != "CREATING":
            break
```

Cette source de données n'a pas de calendrier, elle ne s'exécute donc pas automatiquement. Pour indexer la source de données, vous appelez StartDataSourceSyncJobpour synchroniser l'index avec la source de données.

Les exemples suivants illustrent la synchronisation d'une source de données.

CLI

```
aws kendra start-data-source-sync-job \
 --index-id index ID \
 --id data source ID
```

Python

L'extrait de code Python suivant synchronise une Amazon S3 source de données. Pour un exemple complet, voirDémarrer (AWS SDK pour Python (Boto3)).

```
print("Synchronize the data source.")

    sync_response = kendra.start_data_source_sync_job(
        Id = "data-source-id",
        IndexId = "index-id"
    )
```

Amazon S3 métadonnées du document

Vous pouvez ajouter des métadonnées, des informations supplémentaires sur un document, aux documents d'un Amazon S3 compartiment à l'aide d'un fichier de métadonnées. Chaque fichier de métadonnées est associé à un document indexé.

Vos fichiers de métadonnées doivent être stockés dans le même compartiment que vos fichiers indexés. Vous pouvez spécifier un emplacement dans le compartiment pour vos fichiers de métadonnées à l'aide de la console ou du `S3Prefix` champ du `DocumentsMetadataConfiguration` paramètre lorsque vous créez une source de Amazon S3 données. Si vous ne spécifiez aucun Amazon S3 préfixe, vos fichiers de métadonnées doivent être stockés au même endroit que vos documents indexés.

Si vous spécifiez un Amazon S3 préfixe pour vos fichiers de métadonnées, ceux-ci se trouvent dans une structure de répertoire parallèle à celle de vos documents indexés. Amazon Kendra recherche uniquement dans le répertoire spécifié pour vos métadonnées. Si les métadonnées ne sont pas lues, vérifiez que l'emplacement du répertoire correspond à celui de vos métadonnées.

Les exemples suivants montrent comment l'emplacement du document indexé correspond à l'emplacement du fichier de métadonnées. Notez que la Amazon S3 clé du document est ajoutée au Amazon S3 préfixe des métadonnées, puis suffixée `.metadata.json` pour former le chemin du fichier de métadonnées. Amazon S3 La Amazon S3 clé combinée, avec le Amazon S3 préfixe et le `.metadata.json` suffixe des métadonnées, ne doit pas comporter plus de 1024 caractères au total. Il est recommandé de ne pas dépasser 1 000 caractères pour tenir compte des caractères supplémentaires lorsque vous combinez votre clé avec le préfixe et le suffixe. Amazon S3

```
Bucket name:
```

```
     s3://bucketName
Document path:
     documents
Metadata path:
     none
File mapping
     s3://bucketName/documents/file.txt ->
        s3://bucketName/documents/file.txt.metadata.json
```

```
Bucket name:
     s3://bucketName
Document path:
     documents/legal
Metadata path:
     metadata
File mapping
     s3://bucketName/documents/legal/file.txt ->
        s3://bucketName/metadata/documents/legal/file.txt.metadata.json
```

Les métadonnées de votre document sont définies dans un fichier JSON. Le fichier doit être un fichier texte UTF-8 sans marqueur BOM. Le nom du fichier JSON doit être`<document>.<extension>.metadata.json`. Dans cet exemple, « document » est le nom du document auquel s'appliquent les métadonnées et « extension » est l'extension de fichier du document. L'identifiant du document doit être unique dans`<document>.<extension>.metadata.json`.

Le contenu du fichier JSON suit ce modèle. Tous attributes/fields sont facultatifs, il n'est donc pas nécessaire d'inclure tous les attributs. Vous devez fournir une valeur pour chaque attribut que vous souhaitez inclure ; la valeur ne peut pas être vide. Si vous ne spécifiez pas le`_source_uri`, les liens renvoyés par Amazon Kendra les résultats de recherche pointent vers le Amazon S3 compartiment contenant le document. `DocumentId`est mappé au champ `s3_document_id` et représente le chemin absolu vers le document dans S3.

```
{
    "DocumentId": "S3 document ID, the S3 path to doc",
    "Attributes": {
        "_category": "document category",
        "_created_at": "ISO 8601 encoded string",
        "_last_updated_at": "ISO 8601 encoded string",
        "_source_uri": "document URI",
        "_version": "file version",
```

```
        "_view_count": number of times document has been viewed,
        "custom attribute key": "custom attribute value",
        additional custom attributes
    },
    "AccessControlList": [
         {
             "Name": "user name",
             "Type": "GROUP | USER",
             "Access": "ALLOW | DENY"
         }
    ],
    "Title": "document title",
    "ContentType": "For example HTML | PDF. For supported content types, see Types of
 documents."
}
```

Les champs `_created_at` et `_last_updated_at` les champs de métadonnées sont des dates codées ISO 8601. Par exemple, 2012-03-25T 12:30:10 + 01:00 est le format date-heure ISO 8601 pour le 25 mars 2012 à 12h30 (plus 10 secondes) dans le fuseau horaire d'Europe centrale.

Vous pouvez ajouter des informations supplémentaires au `Attributes` champ concernant un document que vous utilisez pour filtrer les requêtes ou pour regrouper les réponses aux requêtes. Pour de plus amples informations, veuillez consulter Création de champs de document personnalisés.

Vous pouvez utiliser le `AccessControlList` champ pour filtrer la réponse d'une requête. Ainsi, seuls certains utilisateurs et groupes ont accès aux documents. Pour de plus amples informations, veuillez consulter Filtrage en fonction du contexte utilisateur.

Contrôle d'accès pour les sources Amazon S3 de données

Vous pouvez contrôler l'accès aux documents d'une source de Amazon S3 données à l'aide d'un fichier de configuration. Vous spécifiez le fichier dans la console ou en tant que `AccessControlListConfiguration` paramètre lorsque vous appelez l'UpdateDataSourceAPI CreateDataSourceor.

Le fichier de configuration contient une structure JSON qui identifie un préfixe S3 et répertorie les paramètres d'accès pour le préfixe. Le préfixe peut être un chemin ou un fichier individuel. Si le préfixe est un chemin, les paramètres d'accès s'appliquent à tous les fichiers de ce chemin. Le fichier de configuration JSON contient un nombre maximum de préfixes S3 et une taille de fichier maximale par défaut. Pour de plus amples informations, consultez Quotas pour Amazon Kendra.

Vous pouvez spécifier à la fois les utilisateurs et les groupes dans les paramètres d'accès. Lorsque vous interrogez l'index, vous spécifiez les informations relatives aux utilisateurs et aux groupes. Pour de plus amples informations, veuillez consulter Filtrage par attribut utilisateur.

La structure JSON du fichier de configuration doit être au format suivant :

```
[
    {
        "keyPrefix": "s3://BUCKETNAME/prefix1/",
        "aclEntries": [
            {
                "Name": "user1",
                "Type": "USER",
                "Access": "ALLOW"
            },
            {
                "Name": "group1",
                "Type": "GROUP",
                "Access": "DENY"
            }
        ]
    },
    {
        "keyPrefix": "s3://prefix2",
        "aclEntries": [
            {
                "Name": "user2",
                "Type": "USER",
                "Access": "ALLOW"
            },
            {
                "Name": "user1",
                "Type": "USER",
                "Access": "DENY"
            },
            {
                "Name": "group1",
                "Type": "GROUP",
                "Access": "DENY"
            }
        ]
    }
]
```

Utilisation Amazon VPC avec une source Amazon S3 de données

Cette rubrique fournit un step-by-step exemple qui montre comment se connecter à un compartiment Amazon S3 à l'aide d'un connecteur Amazon S3 via Amazon VPC. L'exemple suppose que vous commencez avec un compartiment S3 existant. Nous vous recommandons de ne télécharger que quelques documents dans votre compartiment S3 pour tester l'exemple.

Vous pouvez vous connecter Amazon Kendra à votre Amazon S3 bucket via Amazon VPC. Pour ce faire, vous devez spécifier le Amazon VPC sous-réseau et les groupes de Amazon VPC sécurité lors de la création de votre connecteur de source de Amazon S3 données.

> **Important**
>
> Pour qu'un Amazon Kendra Amazon S3 connecteur puisse accéder à votre Amazon S3 bucket, assurez-vous d'avoir attribué un Amazon S3 point de terminaison à votre cloud privé virtuel (VPC).

Amazon Kendra Pour synchroniser des documents depuis votre Amazon S3 compartiment Amazon VPC, vous devez suivre les étapes suivantes :

- Configurez un Amazon S3 point de terminaison pour Amazon VPC. Pour plus d'informations sur la configuration d'un Amazon S3 point de terminaison, consultez la section Points de terminaison de passerelle pour Amazon S3 le AWS PrivateLink Guide.
- (Facultatif) Vérifiez vos politiques de Amazon S3 bucket pour vous assurer que le Amazon S3 bucket est accessible depuis le cloud privé virtuel (VPC) auquel vous l'avez assigné. Amazon Kendra Pour plus d'informations, consultez la section Contrôle de l'accès depuis les points de terminaison VPC avec des politiques de compartiment dans le guide de l'utilisateur Amazon S3

Étapes

- Étape 1 : Configuration d'un Amazon VPC
- (Facultatif) Étape 2 : Configuration de la politique de Amazon S3 compartiment
- Étape 3 : Création d'un connecteur de source Amazon S3 de données de test

Étape 1 : Configuration d'un Amazon VPC

Créez un réseau VPC comprenant un sous-réseau privé avec un point de terminaison de Amazon S3 passerelle et un groupe de sécurité Amazon Kendra à utiliser ultérieurement.

Pour configurer un VPC avec un sous-réseau privé, un point de terminaison S3 et un groupe de sécurité

1. Connectez-vous à la Amazon VPC console AWS Management Console et ouvrez-la à l'adressehttps://console.aws.amazon.com/vpc/.
2. Créez un VPC avec un sous-réseau privé et un point de terminaison S3 pour Amazon Kendra utiliser :

   Dans le volet de navigation, choisissez Your VPCs, puis Create VPC.

   a. Sous Ressources à créer, choisissez VPC et plus encore.
   b. Pour Name tag, activez Générer automatiquement, puis entrez**`kendra-s3-example`**.
   c. Pour le blocIPv4 / IPv6 CIDR, conservez les valeurs par défaut.
   d. Pour Nombre de zones de disponibilité (AZs), choisissez le numéro 1.
   e. Sélectionnez Personnaliser AZs, puis sélectionnez une zone de disponibilité dans la liste des premières zones de disponibilité.

      Amazon Kendra ne prend en charge qu'un ensemble spécifique de zones de disponibilité.
   f. Pour Nombre de sous-réseaux publics, choisissez le numéro 0.
   g. Pour Nombre de sous-réseaux privés, choisissez le numéro 1.
   h. Pour NAT gateways (Passerelles NAT), choisissez None (Aucun).
   i. Pour les points de terminaison VPC, choisissez Gateway.Amazon S3 .
   j. Conservez les paramètres par défaut pour les autres valeurs.
   k. Sélectionnez Créer VPC.

      Attendez que le flux de travail de création d'un VPC soit terminé. Choisissez ensuite View VPC pour vérifier le VPC que vous venez de créer.

   Vous venez de créer un réseau VPC avec un sous-réseau privé, qui n'a pas accès à l'Internet public.

3. ~~Copiez l'ID de point de terminaison VPC de votre point de terminaison Amazon S3 :~~

a. Dans le volet de navigation, choisissez Points de terminaison.

b. Dans la liste des points de terminaison, recherchez le point de terminaison Amazon S3 `kendra-s3-example-vpce-s3` que vous venez de créer avec votre VPC.

c. Notez l'ID du point de terminaison du VPC.

Vous venez de créer un point de terminaison de passerelle Amazon S3 pour accéder à votre compartiment Amazon S3 via un sous-réseau.

4. Créez un groupe de sécurité Amazon Kendra pour utiliser :

a. Dans le volet de navigation, choisissez Security Groups, puis sélectionnez Create security group.

b. Sous Security group name (Nom du groupe de sécurité), saisissez **`s3-data-source-security-group`**.

c. Choisissez votre VPC dans la Amazon VPCliste.

d. Conservez les règles entrantes et sortantes par défaut.

e. Sélectionnez Create security group (Créer un groupe de sécurité).

Vous venez de créer un groupe de sécurité VPC.

Vous attribuez le sous-réseau et le groupe de sécurité que vous avez créés à votre connecteur de source de données Amazon Kendra Amazon S3 pendant le processus de configuration du connecteur.

(Facultatif) Étape 2 : Configuration de la politique de Amazon S3 compartiment

Dans cette étape facultative, découvrez comment configurer une politique de compartiment Amazon S3 afin que votre compartiment Amazon S3 ne soit accessible que depuis le VPC auquel vous l'attribuez. Amazon Kendra

Amazon Kendra utilise des rôles IAM pour accéder à votre compartiment Amazon S3 et ne nécessite pas que vous configuriez une politique de compartiment Amazon S3. Cependant, il peut être utile de créer une politique de compartiment si vous souhaitez configurer un Amazon S3 connecteur à l'aide d'un compartiment Amazon S3 dont les politiques existantes limitent l'accès depuis l'Internet public.

Pour configurer votre politique de Amazon S3 compartiment

1. Ouvrez la console Amazon S3 à l'adressehttps://console.aws.amazon.com/s3/.
2. Dans le volet de navigation, sélectionnez Buckets.
3. Choisissez le nom du compartiment Amazon S3 avec lequel vous souhaitez effectuer la synchronisation Amazon Kendra.
4. Choisissez l'onglet Autorisations, faites défiler la page jusqu'à Bucket policy, puis cliquez sur Modifier.
5. Ajoutez ou modifiez votre politique de compartiment pour autoriser l'accès uniquement depuis le point de terminaison VPC que vous avez créé.

   Voici un exemple de politique de compartiment. Remplacez *bucket-name* et *vpce-id* par le nom de votre compartiment Amazon S3 et l'ID de point de terminaison Amazon S3 que vous avez indiqués précédemment.

   JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Deny",
      "Principal": "*",
      "Action": "s3:*",
      "Resource": "arn:aws:s3:::bucket-name/*",
      "Condition": {
        "StringNotEquals": {
          "aws:SourceVpce": "vpce-id"
        }
      }
    }
  ]
}
```

6. Sélectionnez Enregistrer les modifications.

Votre compartiment S3 n'est désormais accessible qu'à partir du VPC spécifique que vous avez créé.

Étape 3 : Création d'un connecteur de source Amazon S3 de données de test

Pour tester votre Amazon VPC configuration, créez un Amazon S3 connecteur. Configurez-le ensuite avec le VPC que vous avez créé en suivant les étapes décrites dans. Amazon S3

Pour les valeurs Amazon VPC de configuration, choisissez les valeurs que vous avez créées dans cet exemple :

- Amazon VPC(VPC) — `kendra-s3-example-vpc`
- Sous-réseaux — `kendra-s3-example-subnet-private1-[availability zone]`
- Groupes de sécurité — `s3-data-source-security-group`

Attendez que la création de votre connecteur soit terminée. Une fois le Amazon S3 connecteur créé, choisissez Sync now pour lancer une synchronisation.

La synchronisation peut prendre de quelques minutes à plusieurs heures, selon le nombre de documents contenus dans votre Amazon S3 compartiment. Pour tester cet exemple, nous vous recommandons de ne télécharger que quelques documents dans votre compartiment S3. Si votre configuration est correcte, vous devriez éventuellement voir le statut de synchronisation terminé.

Si vous rencontrez des erreurs, consultez la section Résolution des problèmes de Amazon VPC connexion.

Amazon Kendra Explorateur Web

Vous pouvez utiliser Amazon Kendra Web Crawler pour explorer et indexer des pages Web.

Vous ne pouvez explorer que les sites Web publics ou les sites Web internes de l'entreprise qui utilisent le protocole de communication sécurisé Hypertext Transfer Protocol Secure (HTTPS). Si vous recevez un message d'erreur lors de l'indexation d'un site web, cela signifie peut-être que l'indexation du site web est bloquée. Pour explorer des sites Web internes, vous pouvez configurer un proxy Web. Le proxy Web doit être accessible au public. Vous pouvez également utiliser l'authentification pour accéder à des sites Web et les explorer.

Lorsque vous sélectionnez des sites web à indexer, vous devez respecter les Politiques d'Amazon en matière d'utilisation acceptable et toutes les autres conditions d'Amazon. N'oubliez pas que vous ne devez utiliser Amazon Kendra Web Crawler que pour indexer vos propres pages Web ou les pages Web que vous êtes autorisé à indexer. Pour savoir comment empêcher Amazon Kendra Web Crawler

d'indexer vos sites Web, consultez. Configuration du *robots.txt* fichier pour Amazon Kendra Web Crawler

> **Note**
>
> L'utilisation abusive de Amazon Kendra Web Crawler pour explorer agressivement des sites Web ou des pages Web qui ne vous appartiennent pas n'est pas considérée comme une utilisation acceptable.

Amazon Kendra possède deux versions du web crawler connecteur. Les fonctionnalités prises en charge par chaque version incluent :

Amazon Kendra Connecteur Web Crawler v1.0/WebCrawlerConfigurationAPI

- Proxy Web
- Filtres d'inclusion/exclusion

Amazon Kendra Connecteur Web Crawler v2.0/TemplateConfigurationAPI

- Mappages de champs
- Filtres d'inclusion/exclusion
- Synchronisation complète et incrémentielle du contenu
- Proxy Web
- Authentification de base, NTLM/Kerberos, SAML et par formulaire pour vos sites Web
- Cloud privé virtuel (VPC)

> **Important**
>
> La création de connecteurs Web Crawler v2.0 n'est pas prise en charge par. AWS CloudFormation Utilisez le connecteur Web Crawler v1.0 si vous avez besoin d' AWS CloudFormation assistance.

Pour résoudre les problèmes liés au connecteur de source de données de votre robot d'exploration Amazon Kendra Web, consultezDépannage des sources de données.

Rubriques

- Amazon Kendra Connecteur Web Crawler v1.0
- Amazon Kendra Connecteur Web Crawler v2.0
- Configuration du robots.txt fichier pour Amazon Kendra Web Crawler

Amazon Kendra Connecteur Web Crawler v1.0

Vous pouvez utiliser Amazon Kendra Web Crawler pour explorer et indexer des pages Web.

Vous ne pouvez explorer que les sites Web destinés au public et les sites Web qui utilisent le protocole de communication sécurisé Hypertext Transfer Protocol Secure (HTTPS). Si vous recevez un message d'erreur lors de l'indexation d'un site web, cela signifie peut-être que l'indexation du site web est bloquée. Pour explorer des sites Web internes, vous pouvez configurer un proxy Web. Le proxy Web doit être accessible au public.

Lorsque vous sélectionnez des sites web à indexer, vous devez respecter les Politiques d'Amazon en matière d'utilisation acceptable et toutes les autres conditions d'Amazon. N'oubliez pas que vous ne devez utiliser Amazon Kendra Web Crawler que pour indexer vos propres pages Web ou les pages Web que vous êtes autorisé à indexer. Pour savoir comment empêcher Amazon Kendra Web Crawler d'indexer vos sites Web, consultez. Configuration du *robots.txt* fichier pour Amazon Kendra Web Crawler

> **Note**
>
> L'utilisation abusive de Amazon Kendra Web Crawler pour explorer agressivement des sites Web ou des pages Web qui ne vous appartiennent pas n'est pas considérée comme une utilisation acceptable.

Pour résoudre les problèmes liés au connecteur de source de données de votre robot d'exploration Amazon Kendra Web, consultezDépannage des sources de données.

Rubriques

- Fonctionnalités prises en charge
- Prérequis
- Instructions de connexion

- En savoir plus

Fonctionnalités prises en charge

- Proxy Web
- Filtres d'inclusion/exclusion

Prérequis

Avant de pouvoir Amazon Kendra indexer vos sites Web, vérifiez les détails de vos sites Web et de vos AWS comptes.

Pour vos sites Web, assurez-vous d'avoir :

- Vous avez copié la source ou le plan URLs du site Web que vous souhaitez indexer.
- Pour les sites Web qui nécessitent une authentification de base : notez le nom d'utilisateur et le mot de passe, puis copiez le nom d'hôte du site Web et le numéro de port.
- Facultatif : vous avez copié le nom d'hôte du site Web et le numéro de port si vous souhaitez utiliser un proxy Web pour vous connecter aux sites Web internes que vous souhaitez explorer. Le proxy Web doit être accessible au public. Amazon Kendra prend en charge la connexion à des serveurs proxy Web basés sur une authentification de base ou vous pouvez vous connecter sans authentification.
- Coché : chaque document de page Web que vous souhaitez indexer est unique et que vous comptez utiliser pour le même index parmi les autres sources de données. Chaque source de données que vous souhaitez utiliser pour un index ne doit pas contenir le même document dans toutes les sources de données. IDs Les documents sont globaux par rapport à un index et doivent être uniques par index.

Dans votre AWS compte, assurez-vous d'avoir :

- Création d'un Amazon Kendra index et, si vous utilisez l'API, notez l'ID de l'index.
- Vous avez créé un IAM rôle pour votre source de données et, si vous utilisez l'API, notez l'ARN du IAM rôle.

> **Note**
>
> Si vous modifiez votre type d'authentification et vos informations d'identification, vous devez mettre à jour votre IAM rôle pour accéder au bon identifiant AWS Secrets Manager secret.

- Pour les sites Web qui nécessitent une authentification, ou s'ils utilisent un proxy Web avec authentification, stockez vos informations d'authentification dans un AWS Secrets Manager secret et, si vous utilisez l'API, notez l'ARN du secret.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

Si vous n'avez pas de IAM rôle ou de secret existant, vous pouvez utiliser la console pour créer un nouveau IAM rôle et un nouveau Secrets Manager secret lorsque vous connectez votre web crawler source de données pour Amazon Kendra. Si vous utilisez l'API, vous devez fournir l'ARN d'un IAM rôle et d'un Secrets Manager secret existants, ainsi qu'un identifiant d'index.

Instructions de connexion

Pour vous connecter Amazon Kendra à votre web crawler source de données, vous devez fournir les informations nécessaires sur votre web crawler source de données afin de Amazon Kendra pouvoir accéder à vos données. Si vous n'avez pas encore configuré web crawler pour Amazon Kendra voirPrérequis.

Console

Pour vous connecter Amazon Kendra à web crawler

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la.
2. Dans le volet de navigation de gauche, choisissez Index, puis choisissez l'index que vous souhaitez utiliser dans la liste des index.

> **Note**
>
> Vous pouvez choisir de configurer ou de modifier vos paramètres de contrôle d'accès utilisateur dans les paramètres de l'index.

3. Sur la page de démarrage, choisissez Ajouter une source de données.
4. Sur la page Ajouter une source de données, choisissez le connecteur Web Crawler, puis sélectionnez Ajouter un connecteur. Si vous utilisez la version 2 (le cas échéant), choisissez le connecteur Web Crawler avec la balise « V2.0 ».
5. Sur la page Spécifier les détails de la source de données, entrez les informations suivantes :

   a. Dans Nom et description, pour Nom de la source de données : entrez le nom de votre source de données. Vous pouvez inclure des tirets, mais pas des espaces.

   b. (Facultatif) Description : entrez une description facultative pour votre source de données.

   c. Dans la langue par défaut : choisissez une langue pour filtrer vos documents pour l'index. Sauf indication contraire, la langue par défaut est l'anglais. La langue spécifiée dans les métadonnées du document remplace la langue sélectionnée.

   d. Dans Balises, pour Ajouter une nouvelle balise : incluez des balises facultatives pour rechercher et filtrer vos ressources ou suivre vos AWS coûts.

   e. Choisissez Suivant.
6. Sur la page Définir l'accès et la sécurité, entrez les informations suivantes :

   a. Pour Source, choisissez entre les plans de site Source URLs et Source en fonction de votre cas d'utilisation et entrez les valeurs de chacun.

   Vous pouvez ajouter jusqu'à 10 sources URLs et trois plans de site.

   > **Note**
   >
   > Si vous souhaitez explorer un plan du site, vérifiez que l'URL de base ou racine est identique à celle URLs répertoriée sur votre page de plan du site. Par exemple, si l'URL de votre plan du site est https://example.com/sitemap-page.html, la URLs liste figurant sur cette page de plan du site doit également utiliser l'URL de base »https://example.com/".

   b. (Facultatif) Pour le proxy Web, entrez les informations suivantes :

i. Nom d'hôte : nom d'hôte pour lequel un proxy Web est requis.

ii. Numéro de port : port utilisé par le protocole de transport d'URL de l'hôte. Le numéro de port doit être une valeur numérique comprise entre 0 et 65535.

iii. Pour les informations d'identification du proxy Web : si votre connexion au proxy Web nécessite une authentification, choisissez un secret existant ou créez-en un nouveau pour stocker vos informations d'authentification. Si vous choisissez de créer un nouveau secret, une fenêtre AWS Secrets Manager secrète s'ouvre.

iv. Entrez les informations suivantes dans la fenêtre Créer un AWS Secrets Manager Secrets Manager secret :

A. Nom secret : le nom de votre secret. Le préfixe 'AmazonKendra-WebCrawler-'est automatiquement ajouté à votre nom secret.

B. Pour le nom d'utilisateur et le mot de passe : entrez ces informations d'authentification de base pour vos sites Web.

C. Choisissez Save (Enregistrer).

c. (Facultatif) Hôtes avec authentification : sélectionnez cette option pour ajouter des hôtes supplémentaires avec authentification.

d. IAM rôle —Choisissez un IAM rôle existant ou créez-en un nouveau IAM pour accéder aux informations d'identification de votre référentiel et indexer le contenu.

> **Note**
>
> IAM les rôles utilisés pour les index ne peuvent pas être utilisés pour les sources de données. Si vous ne savez pas si un rôle existant est utilisé pour un index ou une FAQ, choisissez Créer un nouveau rôle pour éviter les erreurs.

e. Choisissez Suivant.

7. Sur la page Configurer les paramètres de synchronisation, entrez les informations suivantes :

a. Plage d'exploration : choisissez le type de pages Web que vous souhaitez explorer.

b. Profondeur du crawl —Sélectionnez le nombre de niveaux à partir de l'URL de départ qui Amazon Kendra doivent être explorés.

c. Paramètres d'exploration avancés et Configuration supplémentaire saisissez les informations suivantes :

i. Taille de fichier maximale : taille maximale de page Web ou de pièce jointe à analyser. Minimum 0,000001 Mo (1 octet). 50 Mo maximum.

ii. Nombre maximum de liens par page : nombre maximal de liens explorés par page. Les liens sont explorés par ordre d'apparition. Minimum 1link/page. Maximum 1000 links/page.

iii. Limitation maximale : nombre maximal d'objets analysés par nom URLs d'hôte et par minute. Minimum 1 URLs /hôtename/minute. Maximum 300 URLs/host name/minute.

iv. Modèles Regex : ajoutez des modèles d'expressions régulières pour inclure ou exclure certains. URLs Vous pouvez ajouter jusqu'à 100 motifs.

d. Dans Synchroniser le calendrier d'exécution, pour Fréquence : choisissez la fréquence de synchronisation avec votre source de données. Amazon Kendra

e. Choisissez Suivant.

8. Sur la page Réviser et créer, vérifiez que les informations saisies sont correctes, puis sélectionnez Ajouter une source de données. Vous pouvez également choisir de modifier vos informations depuis cette page. Votre source de données apparaîtra sur la page Sources de données une fois que la source de données aura été ajoutée avec succès.

API

Pour vous connecter Amazon Kendra à web crawler

Vous devez spécifier les éléments suivants à l'aide de l'WebCrawlerConfigurationAPI :

- URLs—Spécifiez le point de départ ou le point URLs de départ des sites Web ou le plan URLs du site Web que vous souhaitez explorer SeedUrlConfiguration et SiteMapsConfiguration.

> **Note**
>
> Si vous souhaitez explorer un plan du site, vérifiez que l'URL de base ou racine est identique à celle URLs répertoriée sur votre page de plan du site. Par exemple, si l'URL de votre plan du site est https://example.com/sitemap-page.html, la URLs liste figurant sur cette page de plan du site doit également utiliser l'URL de base »https://example.com/".

- Nom de ressource Amazon secret (ARN) : si un site Web nécessite une authentification de base, vous fournissez le nom d'hôte, le numéro de port et un secret qui stocke vos informations d'authentification de base, à savoir votre nom d'utilisateur et votre mot de passe. Vous fournissez l'ARN secret à l'aide du AuthenticationConfigurationAPI. Le secret est stocké dans une structure JSON avec les clés suivantes :

  ```
  {
      "username": "user name",
      "password": "password"
  }
  ```

  Vous pouvez également fournir les informations d'identification du proxy Web à l'aide d'un AWS Secrets Manager secret. Vous utilisez le ProxyConfigurationAPI pour fournir le nom d'hôte et le numéro de port du site Web, et éventuellement le secret qui stocke les informations d'identification de votre proxy Web.
- IAM role —Spécifiez à quel `RoleArn` moment vous appelez `CreateDataSource` pour fournir à un IAM rôle les autorisations d'accéder à votre Secrets Manager secret et pour appeler le public requis APIs pour le connecteur Web Crawler et. Amazon Kendra Pour plus d'informations, consultez la section IAM Rôles des sources de données des robots d'exploration Web.

Vous pouvez également ajouter les fonctionnalités optionnelles suivantes :

- Mode d'exploration : choisissez d'explorer uniquement les noms d'hôtes des sites Web ou les noms d'hôtes avec des sous-domaines, ou d'explorer également les autres domaines vers lesquels les pages Web renvoient.
- La « profondeur » ou le nombre de niveaux entre le niveau de la graine et le rampage. Par exemple, la page URL initiale est de profondeur 1 et tous les hyperliens de cette page qui sont également explorés ont une profondeur de 2.
- Le nombre maximum de pages Web à explorer URLs sur une seule page Web.
- Taille maximale en Mo d'une page Web à explorer.
- Le nombre maximum d' URLs explorations par hôte de site Web par minute.
- L'hôte du proxy Web et le numéro de port permettant de se connecter aux sites Web internes et de les parcourir. Par exemple, le nom d'hôte de https://a.example.com/page1.htmlest »a.example.com« et le numéro de port est 443, le port standard pour HTTPS. Si des informations d'identification de proxy Web sont requises pour se connecter à un hébergeur

de site Web, vous pouvez en créer un AWS Secrets Manager qui stocke les informations d'identification.

- Informations d'authentification permettant d'accéder aux sites Web qui nécessitent une authentification utilisateur et de les analyser.
- Vous pouvez extraire des balises méta HTML sous forme de champs à l'aide de l'outil d'enrichissement de documents personnalisé. Pour plus d'informations, veuillez consulter la rubrique Personnalisation des métadonnées de documents pendant le processus d'intégration. Pour un exemple d'extraction de balises méta HTML, consultez les exemples de CDE.
- Filtres d'inclusion et d'exclusion : spécifiez s'il faut inclure ou exclure certains URLs.

> **Note**
>
> La plupart des sources de données utilisent des modèles d'expressions régulières, qui sont des modèles d'inclusion ou d'exclusion appelés filtres. Si vous spécifiez un filtre d'inclusion, seul le contenu correspondant au filtre d'inclusion est indexé. Tout document qui ne correspond pas au filtre d'inclusion n'est pas indexé. Si vous spécifiez un filtre d'inclusion et d'exclusion, les documents correspondant au filtre d'exclusion ne sont pas indexés, même s'ils correspondent au filtre d'inclusion.

En savoir plus

Pour en savoir plus sur l'intégration Amazon Kendra à votre web crawler source de données, voir :

- Réinventez la découverte des connaissances à l'aide Amazon Kendra du Web Crawler

Amazon Kendra Connecteur Web Crawler v2.0

Vous pouvez utiliser Amazon Kendra Web Crawler pour explorer et indexer des pages Web.

Vous ne pouvez explorer que les sites Web publics ou les sites Web internes de l'entreprise qui utilisent le protocole de communication sécurisé Hypertext Transfer Protocol Secure (HTTPS). Si vous recevez un message d'erreur lors de l'indexation d'un site web, cela signifie peut-être que l'indexation du site web est bloquée. Pour explorer des sites Web internes, vous pouvez configurer un proxy Web. Le proxy Web doit être accessible au public. Vous pouvez également utiliser l'authentification pour accéder à des sites Web et les explorer.

Amazon Kendra Web Crawler v2.0 utilise le package Selenium Web Crawler et un pilote Chromium. Amazon Kendra met automatiquement à jour la version de Selenium et le pilote Chromium à l'aide de l'intégration continue (CI).

Lorsque vous sélectionnez des sites web à indexer, vous devez respecter les Politiques d'Amazon en matière d'utilisation acceptable et toutes les autres conditions d'Amazon. N'oubliez pas que vous ne devez utiliser Amazon Kendra Web Crawler que pour indexer vos propres pages Web ou les pages Web que vous êtes autorisé à indexer. Pour savoir comment empêcher Amazon Kendra Web Crawler d'indexer vos sites Web, consultez. Configuration du *robots.txt* fichier pour Amazon Kendra Web Crawler . L'utilisation abusive de Amazon Kendra Web Crawler pour explorer agressivement des sites Web ou des pages Web qui ne vous appartiennent pas n'est pas considérée comme une utilisation acceptable.

Pour résoudre les problèmes liés au connecteur de source de données de votre robot d'exploration Amazon Kendra Web, consultezDépannage des sources de données.

> **Note**
>
> Le connecteur Web Crawler v2.0 ne prend pas en charge l'analyse de listes de sites Web à partir de AWS KMS compartiments chiffrés. Amazon S3 Il prend uniquement en charge le chiffrement côté serveur avec des clés Amazon S3 gérées.

> **Important**
>
> La création de connecteurs Web Crawler v2.0 n'est pas prise en charge par. AWS CloudFormation Utilisez le connecteur Web Crawler v1.0 si vous avez besoin d' AWS CloudFormation assistance.

Rubriques

- Fonctionnalités prises en charge
- Prérequis
- Instructions de connexion

Fonctionnalités prises en charge

- Mappages de champs

- Filtres d'inclusion/exclusion
- Synchronisation complète et incrémentielle du contenu
- Proxy Web
- Authentification de base, NTLM/Kerberos, SAML et par formulaire pour vos sites Web
- Cloud privé virtuel (VPC)

Prérequis

Avant de pouvoir Amazon Kendra indexer vos sites Web, vérifiez les détails de vos sites Web et de vos AWS comptes.

Pour vos sites Web, assurez-vous d'avoir :

- Vous avez copié la source ou le plan URLs du site Web que vous souhaitez indexer. Vous pouvez les stocker URLs dans un fichier texte et les télécharger dans un Amazon S3 bucket. Chaque URL du fichier texte doit être formatée sur une ligne distincte. Si vous souhaitez stocker vos plans de site dans un Amazon S3 bucket, assurez-vous d'avoir copié le code XML du plan de site et de l'avoir enregistré dans un fichier XML. Vous pouvez également regrouper plusieurs fichiers XML de plan de site dans un fichier ZIP.

  > **Note**
  >
  > (Sur place/sur serveur) Amazon Kendra vérifie si les informations de point de terminaison incluses sont les mêmes AWS Secrets Manager que celles spécifiées dans les détails de configuration de votre source de données. Cela permet de se protéger contre le problème de confusion des adjoints, qui est un problème de sécurité lorsqu'un utilisateur n'est pas autorisé à effectuer une action mais l'utilise Amazon Kendra comme proxy pour accéder au secret configuré et exécuter l'action. Si vous modifiez ultérieurement les informations de votre point de terminaison, vous devez créer un nouveau secret pour synchroniser ces informations.

- Pour les sites Web qui nécessitent une authentification de base, NTLM ou Kerberos :
  - Notez les informations d'authentification de votre site Web, qui incluent un nom d'utilisateur et un mot de passe.

> **Note**
>
> Amazon Kendra Web Crawler v2.0 prend en charge le protocole d'authentification NTLM qui inclut le hachage des mots de passe, et le protocole d'authentification Kerberos qui inclut le chiffrement des mots de passe.

- Pour les sites Web qui nécessitent une authentification SAML ou par formulaire de connexion :
  - Notez les informations d'authentification de votre site Web, qui incluent un nom d'utilisateur et un mot de passe.
  - J'ai copié le XPaths (langage de chemin XML) du champ du nom d'utilisateur (et le bouton du nom d'utilisateur si vous utilisez SAML), du champ de mot de passe et du bouton, et copié l'URL de la page de connexion. Vous pouvez trouver les éléments à l'aide XPaths des outils de développement de votre navigateur Web. XPaths suivent généralement ce format :`//tagname[@Attribute='Value']`.

    > **Note**
    >
    > Amazon Kendra Web Crawler v2.0 utilise un navigateur Chrome sans en-tête et les informations du formulaire pour authentifier et autoriser l'accès avec une OAuth URL protégée 2.0.

- Facultatif : vous avez copié le nom d'hôte et le numéro de port du serveur proxy Web si vous souhaitez utiliser un proxy Web pour vous connecter aux sites Web internes que vous souhaitez explorer. Le proxy Web doit être accessible au public. Amazon Kendra prend en charge la connexion à des serveurs proxy Web basés sur une authentification de base ou vous pouvez vous connecter sans authentification.
- Facultatif : vous avez copié l'ID de sous-réseau du cloud privé virtuel (VPC) si vous souhaitez utiliser un VPC pour vous connecter aux sites Web internes que vous souhaitez explorer. Pour plus d'informations, consultez Configuration d'un Amazon VPC.
- Coché : chaque document de page Web que vous souhaitez indexer est unique et que vous comptez utiliser pour le même index parmi les autres sources de données. Chaque source de données que vous souhaitez utiliser pour un index ne doit pas contenir le même document dans toutes les sources de données. IDs Les documents sont globaux par rapport à un index et doivent être uniques par index.

Dans votre AWS compte, assurez-vous d'avoir :

- Création d'un Amazon Kendra index et, si vous utilisez l'API, notez l'ID de l'index.
- Vous avez créé un IAM rôle pour votre source de données et, si vous utilisez l'API, notez le nom de ressource Amazon du IAM rôle.

> **Note**
>
> Si vous modifiez votre type d'authentification et vos informations d'identification, vous devez mettre à jour votre IAM rôle pour accéder au bon identifiant AWS Secrets Manager secret.

- Pour les sites Web qui nécessitent une authentification, ou s'ils utilisent un proxy Web avec authentification, stockez vos informations d'authentification dans un AWS Secrets Manager secret et, si vous utilisez l'API, notez l'ARN du secret.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

Si vous n'avez pas de IAM rôle ou de secret existant, vous pouvez utiliser la console pour créer un nouveau IAM rôle et un nouveau Secrets Manager secret lorsque vous connectez votre web crawler source de données pour Amazon Kendra. Si vous utilisez l'API, vous devez fournir l'ARN d'un IAM rôle et d'un Secrets Manager secret existants, ainsi qu'un identifiant d'index.

Instructions de connexion

Pour vous connecter Amazon Kendra à votre web crawler source de données, vous devez fournir les informations nécessaires sur votre web crawler source de données afin de Amazon Kendra pouvoir accéder à vos données. Si vous n'avez pas encore configuré web crawler pour Amazon Kendra voirPrérequis.

Console

Pour vous connecter Amazon Kendra à web crawler

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la.
2. Dans le volet de navigation de gauche, choisissez Index, puis choisissez l'index que vous souhaitez utiliser dans la liste des index.

> **Note**
>
> Vous pouvez choisir de configurer ou de modifier vos paramètres de contrôle d'accès utilisateur dans les paramètres de l'index.

3. Sur la page de démarrage, choisissez Ajouter une source de données.
4. Sur la page Ajouter une source de données, choisissez le connecteur Web Crawler, puis sélectionnez Ajouter un connecteur. Si vous utilisez la version 2 (le cas échéant), choisissez le connecteur Web Crawler avec la balise « V2.0 ».
5. Sur la page Spécifier les détails de la source de données, entrez les informations suivantes :

   a. Dans Nom et description, pour Nom de la source de données : entrez le nom de votre source de données. Vous pouvez inclure des tirets, mais pas des espaces.

   b. (Facultatif) Description : entrez une description facultative pour votre source de données.

   c. Dans la langue par défaut : choisissez une langue pour filtrer vos documents pour l'index. Sauf indication contraire, la langue par défaut est l'anglais. La langue spécifiée dans les métadonnées du document remplace la langue sélectionnée.

   d. Dans Balises, pour Ajouter une nouvelle balise : incluez des balises facultatives pour rechercher et filtrer vos ressources ou suivre vos AWS coûts.

   e. Choisissez Suivant.

6. Sur la page Définir l'accès et la sécurité, entrez les informations suivantes :

   a. Source —Choisissez Source, Plans de site source URLs, Fichier source, URLs Fichier de plans de site source. Si vous choisissez d'utiliser un fichier texte contenant une liste contenant jusqu'à 100 graines URLs, vous devez spécifier le chemin d'accès au Amazon S3 compartiment dans lequel votre fichier est stocké. Si vous choisissez d'utiliser un fichier XML de plan de site, vous devez spécifier le chemin d'accès au Amazon S3 compartiment dans lequel votre fichier est stocké. Vous pouvez également regrouper plusieurs fichiers XML de plan de site dans un fichier ZIP. Sinon, vous pouvez saisir

manuellement jusqu'à 10 points de départ ou de départ URLs, et jusqu'à trois plans de site URLs.

> **Note**
>
> Si vous souhaitez explorer un plan du site, vérifiez que l'URL de base ou racine est identique à celle URLs répertoriée sur votre page de plan du site. Par exemple, si l'URL de votre plan du site est https://example.com/sitemap-page.html, la URLs liste figurant sur cette page de plan du site doit également utiliser l'URL de base »https://example.com/".

Si vos sites Web nécessitent une authentification pour accéder aux sites Web, vous pouvez choisir l'authentification ether basic, NTLM/Kerberos, SAML ou par formulaire. Dans le cas contraire, choisissez l'option « Aucune authentification ».

> **Note**
>
> Si vous souhaitez modifier ultérieurement votre source de données pour modifier votre source de données en utilisant URLs l'authentification pour les plans de site, vous devez créer une nouvelle source de données. Amazon Kendra configure la source de données en utilisant les informations du URLs point de terminaison initial contenues dans le Secrets Manager secret à des fins d'authentification, et ne peut donc pas reconfigurer la source de données lors du passage à des plans de site.

- AWS Secrets Manager secret —Si vos sites Web nécessitent la même authentification pour accéder aux sites Web, choisissez un secret existant ou créez-en un nouveau Secrets Manager pour stocker les informations d'identification de votre site Web. Si vous choisissez de créer un nouveau secret, une fenêtre AWS Secrets Manager secrète s'ouvre.

  Si vous avez choisi l'authentification de base ou NTML/Kerberos, entrez le nom du secret, ainsi que le nom d'utilisateur et le mot de passe. Le protocole d'authentification NTLM inclut le hachage des mots de passe et le protocole d'authentification Kerberos inclut le chiffrement des mots de passe.

Si vous avez choisi l'authentification SAML ou par formulaire, entrez le nom du secret, ainsi que le nom d'utilisateur et le mot de passe. XPath À utiliser pour le champ du nom d'utilisateur (et XPath pour le bouton du nom d'utilisateur si vous utilisez SAML). XPaths À utiliser pour le champ et le bouton du mot de passe, ainsi que pour l'URL de la page de connexion. Vous pouvez trouver le XPaths (langage de chemin XML) des éléments à l'aide des outils de développement de votre navigateur Web. XPaths suivent généralement ce format `://tagname[@Attribute='Value']`.

b. (Facultatif) Proxy Web : entrez le nom d'hôte et le numéro de port du serveur proxy que vous souhaitez utiliser pour vous connecter aux sites Web internes. Par exemple, le nom d'hôte de https://a.example.com/page1.htmlest »a.example.com« et le numéro de port est 443, le port standard pour HTTPS. Si des informations d'identification de proxy Web sont requises pour se connecter à un hébergeur de site Web, vous pouvez en créer un AWS Secrets Manager qui stocke les informations d'identification.

c. Virtual Private Cloud (VPC) —Vous pouvez choisir d'utiliser un VPC. Dans ce cas, vous devez ajouter des sous-réseaux et des groupes de sécurité VPC.

d. IAM rôle —Choisissez un IAM rôle existant ou créez-en un nouveau IAM pour accéder aux informations d'identification de votre référentiel et indexer le contenu.

> **Note**
>
> IAM les rôles utilisés pour les index ne peuvent pas être utilisés pour les sources de données. Si vous ne savez pas si un rôle existant est utilisé pour un index ou une FAQ, choisissez Créer un nouveau rôle pour éviter les erreurs.

e. Choisissez Suivant.

7. Sur la page Configurer les paramètres de synchronisation, entrez les informations suivantes :

a. Étendue de synchronisation : définissez des limites pour l'exploration des pages Web, notamment leurs domaines, leurs tailles de fichiers et leurs liens, et filtrez à URLs l'aide de modèles regex.

i. (Facultatif) Plage de domaines d'exploration : choisissez d'explorer uniquement les domaines du site Web, les domaines avec des sous-domaines ou d'explorer également les autres domaines vers lesquels les pages Web renvoient. Par

défaut, explore Amazon Kendra uniquement les domaines des sites Web que vous souhaitez explorer.

ii. (Facultatif) Configuration supplémentaire : définissez les paramètres suivants :

- Profondeur de rampage : « profondeur » ou nombre de niveaux entre le niveau de la graine et le rampage. Par exemple, la page URL initiale est de profondeur 1 et tous les hyperliens de cette page qui sont également explorés ont une profondeur de 2.
- Taille de fichier maximale : taille maximale en Mo d'une page Web ou d'une pièce jointe à analyser.
- Nombre maximum de liens par page : nombre maximum de liens à explorer URLs sur une même page Web.
- Limitation maximale de la vitesse d'exploration : nombre maximal d' URLsexplorations par hôte de site Web et par minute.
- Fichiers : choisissez d'explorer les fichiers vers lesquels les pages Web renvoient.
- Exploration et indexation URLs : ajoutez des modèles d'expressions régulières pour inclure ou exclure l'exploration URLs, certains hyperliens et l'indexation de tous les hyperliens sur ces pages Web URL.

b. Mode de synchronisation : choisissez la manière dont vous souhaitez mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation.

- Synchronisation complète : Indexez tout le contenu fraîchement, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
- Synchronisation nouvelle, modifiée ou supprimée : Indexez uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

c. Synchronisation du calendrier d'exécution : pour Fréquence, choisissez la fréquence à laquelle Amazon Kendra vous souhaitez effectuer la synchronisation avec votre source de données.

d. Choisissez Suivant.

8. Sur la page Définir les mappages de champs, entrez les informations suivantes :

a. Sélectionnez parmi les champs par défaut Amazon Kendra générés des pages Web et des fichiers que vous souhaitez mapper à votre index.

b. Choisissez Suivant.

9. Sur la page Réviser et créer, vérifiez que les informations saisies sont correctes, puis sélectionnez Ajouter une source de données. Vous pouvez également choisir de modifier vos informations depuis cette page. Votre source de données apparaîtra sur la page Sources de données une fois que la source de données aura été ajoutée avec succès.

API

Pour vous connecter Amazon Kendra à web crawler

Vous devez spécifier un code JSON du schéma de source de données à l'aide du TemplateConfigurationAPI. Vous devez fournir les informations suivantes :

- Source de données —Spécifiez le type de source de données `WEBCRAWLERV2` lorsque vous utilisez le TemplateConfigurationSchéma JSON. Spécifiez également la source de données comme `TEMPLATE` lorsque vous appelez le CreateDataSourceAPI.
- URLs—Spécifiez le point de départ ou le point URLs de départ des sites Web ou le plan URLs du site Web que vous souhaitez explorer. Vous pouvez spécifier le chemin d'accès à un Amazon S3 compartiment qui stocke votre liste de graines URLs. Chaque URL du fichier texte pour la source URLs doit être formatée sur une ligne distincte. Vous pouvez également spécifier le chemin d'accès à un Amazon S3 compartiment qui stocke les fichiers XML de votre plan de site. Vous pouvez regrouper plusieurs fichiers de plan du site dans un fichier ZIP et le stocker dans votre Amazon S3 compartiment.

> **Note**
>
> Si vous souhaitez explorer un plan du site, vérifiez que l'URL de base ou racine est identique à celle URLs répertoriée sur votre page de plan du site. Par exemple, si l'URL de votre plan du site est https://example.com/sitemap-page.html, la URLs liste

figurant sur cette page de plan du site doit également utiliser l'URL de base »https://example.com/".

- Mode de synchronisation : spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation. Vous pouvez choisir entre :
  - `FORCED_FULL_CRAWL`pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
  - `FULL_CRAWL`pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
- Authentification : si vos sites Web nécessitent la même authentification, spécifiez soit `BasicAuthNTLM_Kerberos`,`SAML`, soit `Form` authentification. Si vos sites Web ne nécessitent pas d'authentification, spécifiez`NoAuthentication`.
- Nom de ressource Amazon secret (ARN) : si vos sites Web nécessitent une authentification de base, NTLM ou Kerberos, vous fournissez un secret qui stocke vos informations d'authentification, à savoir votre nom d'utilisateur et votre mot de passe. Vous fournissez le Amazon Resource Name (ARN) d'un AWS Secrets Manager secret. Le secret est stocké dans une structure JSON avec les clés suivantes :

```
{
    "seedUrlsHash": "Hash representation of all seed URLs",
    "userName": "user name",
    "password": "password"
}
```

Si vos sites Web nécessitent une authentification SAML, le secret est stocké dans une structure JSON avec les clés suivantes :

```
{
    "seedUrlsHash": "Hash representation of all seed URLs",

    "userName": "user name",
```

```
    "password": "password",
    "userNameFieldXpath": "XPath for user name field",
    "userNameButtonXpath": "XPath for user name button",
    "passwordFieldXpath": "XPath for password field",
    "passwordButtonXpath": "XPath for password button",
    "loginPageUrl": "Full URL for website login page"
}
```

Si vos sites Web nécessitent une authentification par formulaire, le secret est stocké dans une structure JSON avec les clés suivantes :

```
{
    "seedUrlsHash": "Hash representation of all seed URLs",
    "userName": "user name",
    "password": "password",
    "userNameFieldXpath": "XPath for user name field",
    "passwordFieldXpath": "XPath for password field",
    "passwordButtonXpath": "XPath for password button",
    "loginPageUrl": "Full URL for website login page"
}
```

Vous pouvez trouver le XPaths (langage de chemin XML) des éléments à l'aide des outils de développement de votre navigateur Web. XPaths suivent généralement ce format `://tagname[@Attribute='Value']`.

Vous pouvez également fournir les informations d'identification du proxy Web à l'aide d' AWS Secrets Manager un code secret.

- IAM role —Spécifiez à quel `RoleArn` moment vous appelez `CreateDataSource` pour fournir à un IAM rôle les autorisations d'accéder à votre Secrets Manager secret et pour appeler le public requis APIs pour le connecteur Web Crawler et. Amazon Kendra Pour plus d'informations, consultez la section IAM Rôles des sources de données des robots d'exploration Web.

Vous pouvez également ajouter les fonctionnalités optionnelles suivantes :

- Virtual Private Cloud (VPC) **`VpcConfiguration`** —Spécifiez le moment de votre appel. `CreateDataSource` Pour de plus amples informations, veuillez consulter Configuration Amazon Kendra pour utiliser un Amazon VPC.

- Plage de domaines : choisissez d'explorer les domaines des sites Web contenant uniquement des sous-domaines ou d'explorer également les autres domaines vers lesquels les pages Web renvoient. Par défaut, explore Amazon Kendra uniquement les domaines des sites Web que vous souhaitez explorer.
- La « profondeur » ou le nombre de niveaux entre le niveau de la graine et le rampage. Par exemple, la page URL initiale est de profondeur 1 et tous les hyperliens de cette page qui sont également explorés ont une profondeur de 2.
- Le nombre maximum de pages Web à explorer URLs sur une seule page Web.
- Taille maximale en Mo d'une page Web ou d'une pièce jointe à analyser.
- Le nombre maximum d' URLs explorations par hôte de site Web par minute.
- L'hôte du proxy Web et le numéro de port permettant de se connecter aux sites Web internes et de les parcourir. Par exemple, le nom d'hôte de https://a.example.com/page1.htmlest »a.example.com« et le numéro de port est 443, le port standard pour HTTPS. Si des informations d'identification de proxy Web sont requises pour se connecter à un hébergeur de site Web, vous pouvez en créer un AWS Secrets Manager qui stocke les informations d'identification.
- Filtres d'inclusion et d'exclusion : spécifiez s'il faut inclure ou exclure l'exploration de certains hyperliens URLs et l'indexation de tout hyperlien sur ces pages Web URL.

  > **Note**
  >
  > La plupart des sources de données utilisent des modèles d'expressions régulières, qui sont des modèles d'inclusion ou d'exclusion appelés filtres. Si vous spécifiez un filtre d'inclusion, seul le contenu correspondant au filtre d'inclusion est indexé. Tout document qui ne correspond pas au filtre d'inclusion n'est pas indexé. Si vous spécifiez un filtre d'inclusion et d'exclusion, les documents correspondant au filtre d'exclusion ne sont pas indexés, même s'ils correspondent au filtre d'inclusion.

- Mappages de champs : choisissez de mapper les champs des pages Web et des fichiers de pages Web à vos champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.

Pour obtenir la liste des autres clés JSON importantes à configurer, consultez le schéma du modèle Amazon Kendra Web Crawler.

Configuration du **`robots.txt`** fichier pour Amazon Kendra Web Crawler

Amazon Kendra est un service de recherche intelligent que AWS les clients utilisent pour indexer et rechercher les documents de leur choix. Pour indexer des documents sur le Web, les clients peuvent utiliser le Amazon Kendra Web Crawler, qui indique les URL à indexer ainsi que d'autres paramètres opérationnels. Amazon Kendra les clients sont tenus d'obtenir une autorisation avant d'indexer un site Web en particulier.

Amazon Kendra Web Crawler respecte les directives standard de robots.txt telles que `Allow` et`Disallow`. Vous pouvez modifier le `robots.txt` fichier de votre site Web pour contrôler la façon dont Amazon Kendra Web Crawler explore votre site Web.

Configuration de la façon dont Amazon Kendra Web Crawler accède à votre site Web

Vous pouvez contrôler la façon dont le Amazon Kendra Web Crawler indexe votre site Web à l'aide de directives `Allow` et`Disallow`. Vous pouvez également contrôler quelles pages Web sont indexées et quelles pages Web ne sont pas explorées.

Pour autoriser Amazon Kendra Web Crawler à explorer toutes les pages Web à l'exception des pages Web interdites, utilisez la directive suivante :

```
User-agent: amazon-kendra    # Amazon Kendra Web Crawler
Disallow: /credential-pages/ # disallow access to specific pages
```

Pour autoriser Amazon Kendra Web Crawler à explorer uniquement des pages Web spécifiques, utilisez la directive suivante :

```
User-agent: amazon-kendra    # Amazon Kendra Web Crawler
Allow: /pages/ # allow access to specific pages
```

Pour autoriser Amazon Kendra Web Crawler à explorer tout le contenu du site Web et interdire l'exploration à tout autre robot, utilisez la directive suivante :

```
User-agent: amazon-kendra # Amazon Kendra Web Crawler
Allow: / # allow access to all pages
User-agent: * # any (other) robot
Disallow: / # disallow access to any pages
```

Amazon Kendra Empêcher Web Crawler d'explorer votre site Web

Vous pouvez empêcher Amazon Kendra Web Crawler d'indexer votre site Web à l'aide de cette directive. `Disallow` Vous pouvez également contrôler les pages Web qui sont explorées et celles qui ne le sont pas.

Pour empêcher Amazon Kendra Web Crawler d'explorer le site Web, utilisez la directive suivante :

```
User-agent: amazon-kendra # Amazon Kendra Web Crawler
Disallow: / # disallow access to any pages
```

Si vous avez des questions ou des préoccupations concernant Amazon Kendra Web Crawler, vous pouvez contacter l'équipe d'AWS assistance.

Box (Cube)

Box est un service de stockage dans le cloud qui offre des fonctionnalités d'hébergement de fichiers. Vous pouvez l'utiliser Amazon Kendra pour indexer le contenu de votre Box, notamment les commentaires, les tâches et les liens Web.

Vous pouvez vous connecter Amazon Kendra à votre source de données Box à l'aide de la Amazon Kendra console et de l'BoxConfigurationAPI.

Pour résoudre les problèmes liés Amazon Kendra à votre connecteur de source de données Box, consultezDépannage des sources de données.

Rubriques

- Fonctionnalités prises en charge
- Prérequis
- Instructions de connexion
- En savoir plus
- Remarques

Fonctionnalités prises en charge

Amazon Kendra Le connecteur de source de données Box prend en charge les fonctionnalités suivantes :

- Mappages de champs
- Contrôle d'accès des utilisateurs
- Filtres d'inclusion/exclusion
- Journal des modifications, synchronisation complète et incrémentielle du contenu
- Cloud privé virtuel (VPC)

Prérequis

Avant de pouvoir Amazon Kendra indexer votre source de données Box, apportez ces modifications à votre Box et à vos AWS comptes.

Dans Box, assurez-vous d'avoir :

- Un compte Box Enterprise ou Box Enterprise Plus.
- Configuration d'une application personnalisée Box dans la Box Developer Console, avec authentification côté serveur à l'aide de jetons Web JSON (JWT). Consultez la documentation Box sur la création d'une application personnalisée et la documentation Box sur la configuration de JWT Auth pour plus de détails.
- Définissez le niveau d'accès à votre application sur App + Enterprise Access et autorisez-le à effectuer des appels d'API à l'aide de l'en-tête as-user.
- Vous avez utilisé l'utilisateur administrateur pour ajouter les champs d'application suivants dans votre application Box :
  - Écrire tous les fichiers et dossiers stockés dans une boîte
  - Gestion des utilisateurs
  - Gérer les groupes
  - Gérer les propriétés de l'entreprise
- Paire de Public/Private clés configurée comprenant un identifiant client, un secret client, un identifiant de clé publique, un identifiant de clé privée, une phrase secrète et un identifiant d'entreprise à utiliser comme informations d'identification d'authentification. Consultez la section Paire de clés publiques et privées pour plus de détails.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire

pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

- Vous avez copié votre identifiant Box Enterprise depuis les paramètres de votre Box Developer Console ou depuis votre application Box. Par exemple, *801234567*.
- Coché : chaque document est unique dans Box et dans les autres sources de données que vous prévoyez d'utiliser pour le même index. Chaque source de données que vous souhaitez utiliser pour un index ne doit pas contenir le même document dans toutes les sources de données. IDs Les documents sont globaux par rapport à un index et doivent être uniques par index.

Dans votre Compte AWS, assurez-vous d'avoir :

- Création d'un Amazon Kendra index et, si vous utilisez l'API, notez l'ID de l'index.
- Vous avez créé un IAM rôle pour votre source de données et, si vous utilisez l'API, notez l'ARN du IAM rôle.

> **Note**
>
> Si vous modifiez votre type d'authentification et vos informations d'identification, vous devez mettre à jour votre IAM rôle pour accéder au bon identifiant AWS Secrets Manager secret.

- Stockez vos informations d'authentification Box dans un AWS Secrets Manager secret et, si vous utilisez l'API, notez l'ARN du secret.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

Si vous n'avez pas de IAM rôle ou de secret existant, vous pouvez utiliser la console pour créer un nouveau IAM rôle et un nouveau Secrets Manager secret lorsque vous connectez votre source de

données Box à Amazon Kendra. Si vous utilisez l'API, vous devez fournir l'ARN d'un IAM rôle et d'un Secrets Manager secret existants, ainsi qu'un identifiant d'index.

Instructions de connexion

Pour vous connecter Amazon Kendra à votre source de données Box, vous devez fournir les informations nécessaires sur votre source de données Box afin de Amazon Kendra pouvoir accéder à vos données. Si vous n'avez pas encore configuré Box pour Amazon Kendra, consultezPrérequis.

Console

Pour vous connecter Amazon Kendra à Box

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la.
2. Dans le volet de navigation de gauche, choisissez Index, puis choisissez l'index que vous souhaitez utiliser dans la liste des index.

> **Note**
>
> Vous pouvez choisir de configurer ou de modifier vos paramètres de contrôle d'accès utilisateur dans les paramètres de l'index.

3. Sur la page de démarrage, choisissez Ajouter une source de données.
4. Sur la page Ajouter une source de données, choisissez Box connector, puis Add connector. Si vous utilisez la version 2 (le cas échéant), choisissez le connecteur Box avec le tag « V2.0 ».
5. Sur la page Spécifier les détails de la source de données, entrez les informations suivantes :
   a. Dans Nom et description, pour Nom de la source de données : entrez le nom de votre source de données. Vous pouvez inclure des traits d'union, mais pas des espaces.
   b. (Facultatif) Description : entrez une description facultative pour votre source de données.
   c. Dans la langue par défaut : choisissez une langue pour filtrer vos documents pour l'index. Sauf indication contraire, la langue par défaut est l'anglais. La langue spécifiée dans les métadonnées du document remplace la langue sélectionnée.
   d. Dans Balises, pour Ajouter une nouvelle balise : incluez des balises facultatives pour rechercher et filtrer vos ressources ou suivre vos AWS coûts.
   e. Choisissez Suivant.
6. Sur la page Définir l'accès et la sécurité, entrez les informations suivantes :

a. Box Enterprise ID —Entrez votre Box Enterprise ID. Par exemple, *801234567*.

b. Autorisation : activez ou désactivez les informations de la liste de contrôle d'accès (ACL) pour vos documents, si vous disposez d'une ACL et que vous souhaitez l'utiliser pour le contrôle d'accès. L'ACL indique les documents auxquels les utilisateurs et les groupes peuvent accéder. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.

c. AWS Secrets Manager secret —Choisissez un secret existant ou créez-en un nouveau Secrets Manager pour stocker vos informations d'authentification Box. Si vous choisissez de créer un nouveau secret, une fenêtre AWS Secrets Manager secrète s'ouvre.

   i. Nom secret : le nom de votre secret. Le préfixe « AmazonKendra -Box-» est automatiquement ajouté à votre nom secret.

   ii. Pour l'ID client, le secret client, l'ID de clé publique, l'ID de clé privée et le mot de passe, entrez les valeurs de la Public/Private clé que vous avez configurée dans la case.

   iii. Ajoutez et enregistrez votre secret.

d. Virtual Private Cloud (VPC) —Vous pouvez choisir d'utiliser un VPC. Dans ce cas, vous devez ajouter des sous-réseaux et des groupes de sécurité VPC.

e. Identity Crawler : spécifiez s'il faut activer l'explorateur Amazon Kendra d'identité. Le robot d'exploration d'identité utilise les informations de la liste de contrôle d'accès (ACL) de vos documents pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Si vous disposez d'une ACL pour vos documents et que vous choisissez de l'utiliser, vous pouvez également choisir d'activer le robot d'exploration Amazon Kendra d'identité pour configurer le filtrage des résultats de recherche par contexte utilisateur. Sinon, si le robot d'identification est désactivé, tous les documents peuvent être consultés publiquement. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'exploration d'identité est désactivé, vous pouvez également utiliser l'PutPrincipalMappingAPI pour télécharger les informations d'accès des utilisateurs et des groupes afin de filtrer le contexte utilisateur.

f. IAM rôle —Choisissez un IAM rôle existant ou créez-en un nouveau IAM pour accéder aux informations d'identification de votre référentiel et indexer le contenu.

> **Note**
>
> IAM les rôles utilisés pour les index ne peuvent pas être utilisés pour les sources de données. Si vous ne savez pas si un rôle existant est utilisé pour un index ou une FAQ, choisissez Créer un nouveau rôle pour éviter les erreurs.

g. Choisissez Suivant.

7. Sur la page Configurer les paramètres de synchronisation, entrez les informations suivantes :

   a. Fichiers Box : choisissez d'explorer les liens Web, les commentaires et les tâches.

   b. Pour une configuration supplémentaire : ajoutez des modèles d'expressions régulières pour inclure ou exclure certains contenus.

   c. Mode de synchronisation : choisissez la manière dont vous souhaitez mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation.

      - Synchronisation complète : Indexez tout le contenu fraîchement, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
      - Nouvelle synchronisation modifiée : indexez uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
      - Synchronisation nouvelle, modifiée ou supprimée : Indexez uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

   d. Dans Synchroniser le calendrier d'exécution pour la fréquence : choisissez la fréquence à laquelle vous souhaitez synchroniser le contenu de votre source de données et mettre à jour votre index.

e. Choisissez Suivant.

8. Sur la page Définir les mappages de champs, entrez les informations suivantes :

   a. Champs de source de données par défaut : sélectionnez parmi les champs de source de données par défaut Amazon Kendra générés que vous souhaitez mapper à votre index.

   b. Ajouter un champ —Pour ajouter des champs de source de données personnalisés afin de créer un nom de champ d'index à mapper et le type de données du champ.

   c. Choisissez Suivant.

9. Sur la page Réviser et créer, vérifiez que les informations saisies sont correctes, puis sélectionnez Ajouter une source de données. Vous pouvez également choisir de modifier vos informations à partir de cette page. Votre source de données apparaîtra sur la page Sources de données une fois qu'elle aura été ajoutée avec succès.

API

Pour vous connecter Amazon Kendra à Box

Vous devez spécifier les éléments suivants à l'aide de l'BoxConfigurationAPI :

Box Enterprise ID : saisissez votre Box Enterprise ID. Vous trouverez l'identifiant d'entreprise dans les paramètres de la Box Developer Console ou lorsque vous configurez une application dans Box.

- Nom de ressource Amazon (ARN) secret : indiquez le nom de ressource Amazon (ARN) d'un Secrets Manager secret contenant les informations d'authentification de votre compte Box. Le secret est stocké dans une structure JSON avec les clés suivantes :

```
{
    "clientID": "client-id",
    "clientSecret": "client-secret",
    "publicKeyID": "public-key-id",
    "privateKey": "private-key",
    "passphrase": "pass-phrase"
}
```

- IAM role —Spécifiez `RoleArn` quand vous appelez `CreateDataSource` pour fournir à un IAM rôle les autorisations d'accéder à votre Secrets Manager secret et d'appeler le public requis APIs pour le connecteur Box et Amazon Kendra. Pour plus d'informations, consultez la section IAM Rôles des sources de données Box.

Vous pouvez également ajouter les fonctionnalités optionnelles suivantes :

- Virtual Private Cloud (VPC) —Spécifiez dans le `VpcConfiguration` cadre de la configuration de la source de données. Reportez-vous Amazon Kendra à la section Configuration pour utiliser un VPC.
- Journal des modifications : Amazon Kendra faut-il utiliser le mécanisme de journal des modifications de la source de données Box pour déterminer si un document doit être mis à jour dans l'index.

> **Note**
>
> Utilisez le journal des modifications si vous ne Amazon Kendra souhaitez pas numériser tous les documents. Si votre journal des modifications est volumineux, la numérisation des documents de la source de données Box peut prendre Amazon Kendra moins de temps que le traitement du journal des modifications. Si vous synchronisez votre source de données Box avec votre index pour la première fois, tous les documents sont numérisés.

- Commentaires, tâches, liens Web : indiquez si vous souhaitez explorer ces types de contenu.

> **Note**
>
> La plupart des sources de données utilisent des modèles d'expressions régulières, qui sont des modèles d'inclusion ou d'exclusion appelés filtres. Si vous spécifiez un filtre d'inclusion, seul le contenu correspondant au filtre d'inclusion est indexé. Tout document qui ne correspond pas au filtre d'inclusion n'est pas indexé. Si vous spécifiez un filtre d'inclusion et d'exclusion, les documents correspondant au filtre d'exclusion ne sont pas indexés, même s'ils correspondent au filtre d'inclusion.

- Filtres d'inclusion et d'exclusion : indiquez si vous souhaitez inclure ou exclure certains fichiers et dossiers Box.

> **Note**
>
> La plupart des sources de données utilisent des modèles d'expressions régulières, qui sont des modèles d'inclusion ou d'exclusion appelés filtres. Si vous spécifiez un filtre d'inclusion, seul le contenu correspondant au filtre d'inclusion est indexé. Tout document qui ne correspond pas au filtre d'inclusion n'est pas indexé. Si vous spécifiez

un filtre d'inclusion et d'exclusion, les documents correspondant au filtre d'exclusion ne sont pas indexés, même s'ils correspondent au filtre d'inclusion.

- Filtrage du contexte utilisateur et contrôle d'accèsAmazon Kendra : analyse la liste de contrôle d'accès (ACL) de vos documents, si vous disposez d'une ACL pour vos documents. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.
- Mappages de champs : choisissez de mapper les champs de votre source de données Box à vos champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.

> **Note**
>
> Le champ du corps du document ou l'équivalent du corps du document pour vos documents est requis pour Amazon Kendra effectuer une recherche dans vos documents. Vous devez associer le nom du champ du corps du document dans votre source de données au nom du champ d'index`_document_body`. Tous les autres champs sont facultatifs.

En savoir plus

Pour en savoir plus sur l'intégration Amazon Kendra à votre source de données Box, consultez :

- Commencer à utiliser le connecteur Amazon Kendra Box

Remarques

- Lorsque les listes de contrôle d'accès (ACLs) sont activées, l'option « Synchroniser uniquement le contenu nouveau ou modifié » n'est pas disponible en raison des limites de l'API Box. Nous vous recommandons d'utiliser plutôt les modes « Synchronisation complète » ou « Synchronisation du contenu nouveau, modifié ou supprimé », ou de les désactiver ACLs si vous devez utiliser ce mode de synchronisation.

Confluence

Confluence est un outil de gestion du travail collaboratif conçu pour partager, stocker et travailler sur la planification de projets, le développement de logiciels et la gestion de produits. Amazon Kendra prend en charge à la fois Confluence Server/Data Center et Confluence Cloud. Vous pouvez les utiliser Amazon Kendra pour indexer les entités Confluence suivantes :

- Espaces : zones désignées de haut niveau pour organiser le contenu connexe. Chaque espace sert de conteneur, capable de contenir plusieurs pages, blogs et pièces jointes.
- Pages : documents individuels dans un espace où les utilisateurs créent et gèrent du contenu. Les pages peuvent contenir du texte, des images, des tableaux et des éléments multimédias, et peuvent comporter des sous-pages imbriquées. Chaque page est considérée comme un document unique.
- Blogs : contenu similaire à des pages, généralement utilisé pour des mises à jour ou des annonces. Chaque article de blog est considéré comme un document unique.
- Commentaires — Permet aux utilisateurs de donner leur avis ou de participer à des discussions sur un contenu spécifique dans des pages ou des articles de blog.
- Pièces jointes : fichiers chargés sur des pages ou des articles de blog dans Confluence, tels que des images, des documents ou d'autres types de fichiers.

Par défaut, Amazon Kendra n'indexe pas les archives Confluence ni les espaces personnels. Vous pouvez choisir de les indexer lorsque vous créez la source de données. Si vous ne souhaitez pas Amazon Kendra indexer un espace, marquez-le comme privé dans Confluence.

Vous pouvez vous connecter Amazon Kendra à votre source de données Confluence à l'aide de la Amazon Kendra console, de l'TemplateConfigurationAPI ou de l'ConfluenceConfigurationAPI.

Amazon Kendra possède deux versions du connecteur Confluence. Les fonctionnalités suivantes sont prises en charge.

Connecteur Confluence V2.0/API TemplateConfiguration

- Mappages de champs
- Contrôle d'accès utilisateur
- Modèles d'inclusion/exclusion
- Synchronisation complète et incrémentielle du contenu
- Cloud privé virtuel (VPC)

Connecteur Confluence V1.0/ConfluenceConfigurationAPI (n'est plus pris en charge)

- Mappages de champs
- Contrôle d'accès utilisateur
- Filtres d'inclusion/exclusion
- (Confluence Server uniquement) Cloud privé virtuel (VPC)

> **Note**
>
> Le connecteur Confluence ConfluenceConfiguration V1.0/API a pris fin en 2023. Nous vous recommandons de migrer vers ou d'utiliser le connecteur Confluence V2.0/API. TemplateConfiguration

Pour résoudre les problèmes liés à votre connecteur de source de données Amazon Kendra Confluence, consultezDépannage des sources de données.

Rubriques

- ACLs dans Confluence Connector
- Connecteur Confluence V2.0
- Connecteur Confluence V1.0

ACLs dans Confluence Connector

Les connecteurs permettent d'analyser les listes de contrôle d'accès (ACLs) et d'identifier les informations, le cas échéant, en fonction de la source de données. Si vous indexez des documents sans ACLs, tous les documents sont considérés comme publics. L'indexation des documents ACLs garantit la sécurité des données.

Le connecteur Amazon Kendra Confluence scanne les espaces pour collecter les pages et les articles de blog ainsi que leurs ACLs informations. Si aucune restriction n'est appliquée à une page ou à un blog, le connecteur hérite des autorisations de son espace. Si une restriction d'utilisateur ou de groupe spécifique est appliquée sur une page, seuls ces utilisateurs pourront accéder à cette page. Si la page est imbriquée, elle hérite des autorisations de la page parent si aucune restriction n'est appliquée. Un modèle d'autorisation similaire s'applique aux blogs ; toutefois, Confluence ne prend pas en charge les blogs imbriqués.

En outre, le connecteur Amazon Kendra Confluence analyse les informations principales de l'utilisateur (alias d'utilisateur local, configurations d'identité de groupe local et de groupe fédéré) et ses autorisations pour chaque espace configuré.

> **Note**
>
> Le connecteur Confluence Cloud ne prend pas en charge les macros d'exploration, les tableaux blancs ou les bases de données.

Le connecteur Amazon Kendra Confluence met à jour les modifications de l'ACL chaque fois qu'il analyse le contenu de votre source de données. Pour vous assurer que les bons utilisateurs ont accès au contenu approprié, resynchronisez régulièrement votre source de données pour capturer les mises à jour de l'ACL.

Connecteur Confluence V2.0

Confluence est un outil de gestion du travail collaboratif conçu pour partager, stocker et travailler sur la planification de projets, le développement de logiciels et la gestion de produits. Vous pouvez l'utiliser Amazon Kendra pour indexer vos espaces Confluence, vos pages (y compris les pages imbriquées), vos blogs, ainsi que vos commentaires et pièces jointes vers des pages et des blogs indexés.

Pour résoudre les problèmes liés à votre connecteur de source de données Amazon Kendra Confluence, consultezDépannage des sources de données.

Rubriques

- Fonctionnalités prises en charge
- Prérequis
- Instructions de connexion

Fonctionnalités prises en charge

Amazon Kendra Le connecteur de source de données Confluence prend en charge les fonctionnalités suivantes :

- Mappages de champs
- Contrôle d'accès utilisateur

- Modèles d'inclusion/exclusion
- Synchronisation complète et incrémentielle du contenu
- Cloud privé virtuel (VPC)

Prérequis

Avant de pouvoir Amazon Kendra indexer votre source de données Confluence, apportez ces modifications à votre Confluence et AWS à vos comptes.

Dans Confluence, assurez-vous que vous disposez des éléments suivants :

- Vous avez copié l'URL de votre instance Confluence. Par exemple :*https://example.confluence.com*, ou*https://www.example.confluence.com/*, ou*https://atlassian.net/*. Vous avez besoin de l'URL de votre instance Confluence pour vous connecter Amazon Kendra.

  Si vous utilisez Confluence Cloud, l'URL de votre hôte doit se terminer par*atlassian.net/*.

> **Note**
>
> Les formats d'URL suivants ne sont pas pris en charge :
>
> - *https://example.confluence.com/xyz*
> - *https://www.example.confluence.com//wiki/spacekey/xxx*
> - *https://atlassian.net/xyz*

> **Note**
>
> (Sur place/sur serveur) Amazon Kendra vérifie si les informations de point de terminaison incluses sont les mêmes AWS Secrets Manager que celles spécifiées dans les détails de configuration de votre source de données. Cela permet de se protéger contre le problème de confusion des adjoints, qui est un problème de sécurité lorsqu'un utilisateur n'est pas autorisé à effectuer une action mais l'utilise Amazon Kendra comme proxy pour accéder au secret configuré et exécuter l'action. Si vous modifiez ultérieurement les informations de votre point de terminaison, vous devez créer un nouveau secret pour synchroniser ces informations.

- Identifiants d'authentification de base configurés contenant un nom d'utilisateur (identifiant e-mail utilisé pour se connecter à Confluence) et un mot de passe (jeton d'API Confluence comme mot de passe). Consultez Gérer les jetons d'API pour votre compte Atlassian.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

- Facultatif : informations d'identification OAuth 2.0 configurées contenant une clé d'application Confluence, un secret d'application Confluence, un jeton d'accès Confluence et un jeton d'actualisation Confluence pour permettre la connexion Amazon Kendra à votre instance Confluence. Si votre jeton d'accès expire, vous pouvez utiliser le jeton d'actualisation pour régénérer votre jeton d'accès et actualiser la paire de jetons. Vous pouvez également répéter le processus d'autorisation. Pour plus d'informations sur les jetons d'accès, voir Gérer les jetons OAuth d'accès.
- (Pour le serveur ou le centre de données Confluence uniquement) Facultatif : vous avez configuré un jeton d'accès personnel (PAT) dans Confluence. Consultez la section Utilisation de jetons d'accès personnels.

Dans votre Compte AWS, assurez-vous d'avoir :

- Création d'un Amazon Kendra index et, si vous utilisez l'API, notez l'ID de l'index.
- Vous avez créé un IAM rôle pour votre source de données et, si vous utilisez l'API, notez l'ARN du IAM rôle.

> **Note**
>
> Si vous modifiez votre type d'authentification et vos informations d'identification, vous devez mettre à jour votre IAM rôle pour accéder au bon identifiant AWS Secrets Manager secret.

- Stockez vos informations d'authentification Confluence dans un AWS Secrets Manager secret et, si vous utilisez l'API, notez l'ARN du secret.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

Si vous n'avez pas de IAM rôle ou de secret existant, vous pouvez utiliser la console pour créer un nouveau IAM rôle et un nouveau Secrets Manager secret lorsque vous connectez votre source de données Confluence à Amazon Kendra. Si vous utilisez l'API, vous devez fournir l'ARN d'un IAM rôle et d'un Secrets Manager secret existants, ainsi qu'un identifiant d'index.

Instructions de connexion

Pour vous connecter Amazon Kendra à votre source de données Confluence, vous devez fournir les informations nécessaires sur votre source de données Confluence afin de Amazon Kendra pouvoir accéder à vos données. Si vous n'avez pas encore configuré Confluence pour Amazon Kendra voirPrérequis.

Console

Pour vous connecter Amazon Kendra à Confluence

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la.
2. Dans le volet de navigation de gauche, choisissez Index, puis choisissez l'index que vous souhaitez utiliser dans la liste des index.

   > **Note**
   >
   > Vous pouvez choisir de configurer ou de modifier vos paramètres de contrôle d'accès utilisateur dans les paramètres de l'index.

3. Sur la page de démarrage, choisissez Ajouter une source de données.
4. Sur la page Ajouter une source de données, choisissez le connecteur Confluence, puis sélectionnez Ajouter un connecteur. Si vous utilisez la version 2 (le cas échéant), choisissez le connecteur Confluence avec le tag « V2.0 ».

5. Sur la page Spécifier les détails de la source de données, entrez les informations suivantes :

   a. Dans Nom et description, pour Nom de la source de données : entrez le nom de votre source de données. Vous pouvez inclure des tirets, mais pas des espaces.

   b. (Facultatif) Description : entrez une description facultative pour votre source de données.

   c. Dans la langue par défaut : choisissez une langue pour filtrer vos documents pour l'index. Sauf indication contraire, la langue par défaut est l'anglais. La langue spécifiée dans les métadonnées du document remplace la langue sélectionnée.

   d. Dans Balises, pour Ajouter une nouvelle balise : incluez des balises facultatives pour rechercher et filtrer vos ressources ou suivre vos AWS coûts.

   e. Choisissez Suivant.

6. Sur la page Définir l'accès et la sécurité, entrez les informations suivantes :

   a. Dans Source, choisissez Confluence Cloud ou Confluence Server/Data Center.

   b. URL de confluence —Entrez l'URL de l'hôte Confluence. Par exemple, *`https://example.confluence.com`*.

   c. (Pour Confluence Server/Data Center uniquement) Emplacement du certificat SSL - facultatif —Entrez le Amazon S3 chemin d'accès à votre fichier de certificat SSL pour Confluence Server.

   d. (Pour le serveur Confluence ou le centre de données uniquement) Proxy Web - facultatif : entrez le nom d'hôte du proxy Web (sans le `https://` protocole `http://` or) et le numéro de port (port utilisé par le protocole de transport d'URL de l'hôte). Le numéro de port doit être une valeur numérique comprise entre 0 et 65535.

   e. Autorisation : activez ou désactivez les informations de la liste de contrôle d'accès (ACL) pour vos documents, si vous disposez d'une ACL et que vous souhaitez l'utiliser pour le contrôle d'accès. L'ACL indique les documents auxquels les utilisateurs et les groupes peuvent accéder. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.

   f. Authentification : choisissez l'authentification de base, l'authentification Oauth 2.0 ou (pour Confluence Server/Data Center uniquement) l'authentification par jeton d'accès personnel.

   g. AWS Secrets Manager secret —Choisissez un secret existant ou créez-en un nouveau Secrets Manager pour stocker vos identifiants d'authentification Confluence. Si vous

choisissez de créer un nouveau secret, une fenêtre AWS Secrets Manager secrète s'ouvre. Entrez les informations suivantes dans la fenêtre :

i. Nom secret : le nom de votre secret. Le préfixe « AmazonKendra -Confluence-» est automatiquement ajouté à votre nom secret.

ii. Si vous utilisez l'authentification de base, entrez le nom secret, le nom d'utilisateur et le mot de passe (jeton d'API Confluence comme mot de passe) que vous avez configurés dans Confluence.

Si vous utilisez l'authentification OAuth2 .0 : entrez le nom secret, la clé de l'application, le secret de l'application, le jeton d'accès et le jeton d'actualisation que vous avez configurés dans Confluence.

(Serveur/centre de données Confluence uniquement) Si vous utilisez l'authentification par jeton d'accès personnel, entrez le nom secret et le jeton Confluence que vous avez configurés dans votre Confluence.

iii. Enregistrez et ajoutez votre secret.

h. Virtual Private Cloud (VPC) —Vous pouvez choisir d'utiliser un VPC. Dans ce cas, vous devez ajouter des sous-réseaux et des groupes de sécurité VPC.

i. Identity Crawler : spécifiez s'il faut activer l'explorateur Amazon Kendra d'identité. Le robot d'exploration d'identité utilise les informations de la liste de contrôle d'accès (ACL) de vos documents pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Si vous disposez d'une ACL pour vos documents et que vous choisissez de l'utiliser, vous pouvez également choisir d'activer le robot d'exploration Amazon Kendra d'identité pour configurer le filtrage des résultats de recherche par contexte utilisateur. Sinon, si le robot d'identification est désactivé, tous les documents peuvent être consultés publiquement. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'exploration d'identité est désactivé, vous pouvez également utiliser l'PutPrincipalMappingAPI pour télécharger les informations d'accès des utilisateurs et des groupes afin de filtrer le contexte utilisateur.

j. IAM rôle —Choisissez un IAM rôle existant ou créez-en un nouveau IAM pour accéder aux informations d'identification de votre référentiel et indexer le contenu.

> **Note**
>
> IAM les rôles utilisés pour les index ne peuvent pas être utilisés pour les sources de données. Si vous ne savez pas si un rôle existant est utilisé pour un index ou une FAQ, choisissez Créer un nouveau rôle pour éviter les erreurs.

k. Choisissez Suivant.

7. Sur la page Configurer les paramètres de synchronisation, entrez les informations suivantes :

   a. Dans la zone de synchronisation, pour Synchroniser le contenu : choisissez de synchroniser les types de contenu suivants : pages, commentaires de page, pièces jointes de page, blogs, commentaires de blog, pièces jointes de blog, espaces personnels et espaces archivés.

   > **Note**
   >
   > Les commentaires de page et les pièces jointes aux pages ne peuvent être sélectionnés que si vous choisissez de synchroniser les pages. Les commentaires de blog et les pièces jointes de blog ne peuvent être sélectionnés que si vous choisissez de synchroniser les blogs.

   > **Important**
   >
   > Si vous ne spécifiez pas de modèle d'expression régulière avec touche d'espace dans Configuration supplémentaire, toutes les pages et tous les blogs seront explorés par défaut.

   b. Dans Configuration supplémentaire, pour Taille de fichier maximale, spécifiez la limite de taille de fichier MBs à Amazon Kendra analyser. Amazon Kendra explorera uniquement les fichiers dans la limite de taille que vous avez définie. La taille de fichier par défaut est de 50 Mo. La taille maximale du fichier doit être supérieure à 0 Mo et inférieure ou égale à 50 Mo.

   Pour les modèles de regex Spaces : spécifiez si vous souhaitez inclure ou exclure des espaces spécifiques dans votre index en utilisant :

- Touche espace (par exemple,*my-space-123*)

> **Note**
>
> Si vous ne spécifiez pas de modèle d'expression régulière avec une touche espace, toutes les pages et tous les blogs seront explorés par défaut.

- URL (par exemple,*.*/MySite/MyDocuments/*)
- Type de fichier (par exemple,*.*\.pdf, .*\.txt*)

Pour les modèles d'expression régulière des titres d'entités : spécifiez des modèles d'expressions régulières pour inclure ou exclure certains blogs, pages, commentaires et pièces jointes par titre.

> **Note**
>
> Si vous souhaitez inclure ou exclure l'exploration d'une page ou d'une sous-page spécifique, vous pouvez utiliser des modèles réguliers de titre de page.

c. Mode de synchronisation : choisissez la manière dont vous souhaitez mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation.

- Synchronisation complète : Indexez tout le contenu fraîchement, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
- Synchronisation nouvelle, modifiée ou supprimée : Indexez uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

d. Dans Calendrier d'exécution de la synchronisation, pour Fréquence : choisissez la fréquence à laquelle vous souhaitez synchroniser le contenu de votre source de données et mettre à jour votre index.

e. Choisissez Suivant.

8. Sur la page Définir les mappages de champs, entrez les informations suivantes :

   a. Sélectionnez parmi les champs de source de données par défaut Amazon Kendra générés que vous souhaitez mapper à votre index. Pour ajouter des champs de source de données personnalisés, créez un nom de champ d'index à mapper et le type de données du champ.

   b. Choisissez Suivant.

9. Sur la page Réviser et créer, vérifiez que les informations saisies sont correctes, puis sélectionnez Ajouter une source de données. Vous pouvez également choisir de modifier vos informations depuis cette page. Votre source de données apparaîtra sur la page Sources de données une fois que la source de données aura été ajoutée avec succès.

API

Pour vous connecter Amazon Kendra à Confluence

Vous devez spécifier un JSON du schéma de source de données à l'aide de l'TemplateConfigurationAPI. Vous devez fournir les informations suivantes :

- Source de données —Spécifiez le type de source de données `CONFLUENCEV2` lorsque vous utilisez le TemplateConfigurationSchéma JSON. Spécifiez également la source de données comme `TEMPLATE` lorsque vous appelez le CreateDataSourceAPI.
- URL de l'hôte —Spécifiez l'instance de l'URL hôte Confluence. Par exemple, *`https://example.confluence.com`*.
- Mode de synchronisation : spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation. Vous pouvez choisir entre :
  - `FORCED_FULL_CRAWL`pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.

  - `FULL_CRAWL`pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
- Type d'authentification : spécifiez le type d'authentification, si`Basic`,`0Auth2`, (Confluence Server uniquement). `Personal-token`
- (Facultatif, pour Confluence Server uniquement) Emplacement du certificat SSL : spécifiez le `S3bucketName` et que `s3certificateName` vous avez utilisé pour stocker votre certificat SSL.
- Nom de ressource Amazon (ARN) secret : indiquez le nom de ressource Amazon (ARN) d'un Secrets Manager secret contenant les informations d'authentification que vous avez configurées dans Confluence. Si vous utilisez l'authentification de base, le secret est stocké dans une structure JSON avec les clés suivantes :

```
{
    "username": "email ID or user name",
    "password": "Confluence API token"
}
```

Si vous utilisez l'authentification OAuth 2.0, le secret est stocké dans une structure JSON avec les clés suivantes :

```
{
    "confluenceAppKey": "app key",
    "confluenceAppSecret": "app secret",
    "confluenceAccessToken": "access token",
    "confluenceRefreshToken": "refresh token"
}
```

(Pour Confluence Server uniquement) Si vous utilisez l'authentification de base, le secret est stocké dans une structure JSON avec les clés suivantes :

```
{
    "hostUrl": "Confluence Server host URL",
    "username": "Confluence Server user name",
    "password": "Confluence Server password"
}
```

(Pour Confluence Server uniquement) Si vous utilisez l'authentification par jeton d'accès personnel, le secret est stocké dans une structure JSON avec les clés suivantes :

```
{
    "hostUrl": "Confluence Server host URL",
    "patToken": "personal access token"
}
```

- IAM role —Spécifiez à quel `RoleArn` moment vous appelez `CreateDataSource` pour fournir à un IAM rôle les autorisations d'accéder à votre Secrets Manager secret et d'appeler le public requis APIs pour le connecteur Confluence et. Amazon Kendra Pour plus d'informations, consultez la section IAM Rôles des sources de données Confluence.

Vous pouvez également ajouter les fonctionnalités optionnelles suivantes :

- Virtual Private Cloud (VPC) **`VpcConfiguration`** —Spécifiez le moment de votre appel. `CreateDataSource` Pour de plus amples informations, veuillez consulter Configuration Amazon Kendra pour utiliser un Amazon VPC.
- Taille du fichier —Spécifiez la taille maximale du fichier à analyser.
- Types de documents/contenus : indiquez si vous souhaitez explorer les pages, les commentaires de page, les pièces jointes aux pages, les blogs, les commentaires de blog, les pièces jointes de blog, les espaces et les espaces archivés.
- Filtres d'inclusion et d'exclusion : indiquez s'il faut inclure ou exclure certains espaces, pages, blogs, ainsi que leurs commentaires et pièces jointes.

> **Note**
>
> La plupart des sources de données utilisent des modèles d'expressions régulières, qui sont des modèles d'inclusion ou d'exclusion appelés filtres. Si vous spécifiez un filtre d'inclusion, seul le contenu correspondant au filtre d'inclusion est indexé. Tout document qui ne correspond pas au filtre d'inclusion n'est pas indexé. Si vous spécifiez un filtre d'inclusion et d'exclusion, les documents correspondant au filtre d'exclusion ne sont pas indexés, même s'ils correspondent au filtre d'inclusion.

- Proxy Web —Spécifiez les informations de votre proxy Web si vous souhaitez vous connecter à votre instance d'URL Confluence via un proxy Web. Vous pouvez utiliser cette option pour Confluence Server.

- Liste de contrôle d'accès (ACL) : indiquez si vous souhaitez analyser les informations ACL de vos documents, si vous disposez d'une ACL et souhaitez l'utiliser pour le contrôle d'accès. L'ACL indique les documents auxquels les utilisateurs et les groupes peuvent accéder. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.
- Identity Crawler : spécifiez s'il faut activer l'explorateur Amazon Kendra d'identité. Le robot d'exploration d'identité utilise les informations de la liste de contrôle d'accès (ACL) de vos documents pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Si vous disposez d'une ACL pour vos documents et que vous choisissez de l'utiliser, vous pouvez également choisir d'activer le robot d'exploration Amazon Kendra d'identité pour configurer le filtrage des résultats de recherche par contexte utilisateur. Sinon, si le robot d'identification est désactivé, tous les documents peuvent être consultés publiquement. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'exploration d'identité est désactivé, vous pouvez également utiliser l'PutPrincipalMappingAPI pour télécharger les informations d'accès des utilisateurs et des groupes afin de filtrer le contexte utilisateur.
- Mappages de champs : choisissez de mapper les champs de votre source de données Confluence à vos Amazon Kendra champs d'index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.

  > **Note**
  >
  > Le champ du corps du document ou l'équivalent du corps du document pour vos documents est requis pour Amazon Kendra effectuer une recherche dans vos documents. Vous devez associer le nom du champ du corps du document dans votre source de données au nom du champ d'index`_document_body`. Tous les autres champs sont facultatifs.

Pour obtenir la liste des autres clés JSON importantes à configurer, consultez le schéma du modèle Confluence.

Remarques

- Le jeton d'accès personnel (PAT) n'est pas disponible pour Confluence Cloud.

Connecteur Confluence V1.0

Confluence est un outil de gestion du travail collaboratif conçu pour partager, stocker et travailler sur la planification de projets, le développement de logiciels et la gestion de produits. Vous pouvez l'utiliser Amazon Kendra pour indexer vos espaces Confluence, vos pages (y compris les pages imbriquées), vos blogs, ainsi que vos commentaires et pièces jointes vers des pages et des blogs indexés.

> **Note**
>
> Le connecteur Confluence ConfluenceConfiguration V1.0/API a pris fin en 2023. Nous vous recommandons de migrer vers ou d'utiliser le connecteur Confluence V2.0/API. TemplateConfiguration

Pour résoudre les problèmes liés à votre connecteur de source de données Amazon Kendra Confluence, consultezDépannage des sources de données.

Rubriques

- Fonctionnalités prises en charge
- Prérequis
- Instructions de connexion
- En savoir plus

Fonctionnalités prises en charge

Amazon Kendra Le connecteur de source de données Confluence prend en charge les fonctionnalités suivantes :

- Mappages de champs
- Contrôle d'accès utilisateur
- Filtres d'inclusion/exclusion
- (Pour Confluence Server uniquement) Cloud privé virtuel (VPC)

Prérequis

Avant de pouvoir Amazon Kendra indexer votre source de données Confluence, apportez ces modifications à votre Confluence et AWS à vos comptes.

Dans Confluence, assurez-vous que vous disposez des éléments suivants :

- Amazon Kendra Autorisations accordées pour afficher tout le contenu de votre instance Confluence en :
  - Amazon Kendra Devenir membre d'un `confluence-administrators` groupe.
  - Accorder des autorisations d'administrateur de site pour tous les espaces, blogs et pages existants.
- Vous avez copié l'URL de votre instance Confluence.
- Pour les utilisateurs du SSO (Single Sign-On) : activation de la page Afficher lors de la connexion pour le nom d'utilisateur et le mot de passe lorsque vous configurez les méthodes d'authentification Confluence dans Confluence Data Center.
- Pour Confluence Server
  - Vous avez noté vos informations d'authentification de base contenant le nom d'utilisateur et le mot de passe de votre compte administratif Confluence auxquels vous pouvez vous connecter Amazon Kendra.

    > **Note**
    >
    > Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

  - Facultatif : vous avez généré un jeton d'accès personnel dans votre compte Confluence auquel vous connecter Amazon Kendra. Pour plus d'informations, consultez la documentation de Confluence sur la génération de jetons d'accès personnels.
- Pour Confluence Cloud
  - Vous avez noté vos informations d'authentification de base contenant le nom d'utilisateur et le mot de passe de votre compte administratif Confluence auxquels vous pouvez vous connecter Amazon Kendra.

- Il est vérifié que chaque document est unique dans Confluence et dans les autres sources de données que vous prévoyez d'utiliser pour le même index. Chaque source de données que vous souhaitez utiliser pour un index ne doit pas contenir le même document dans toutes les sources de données. IDs Les documents sont globaux par rapport à un index et doivent être uniques par index.

Dans votre Compte AWS, assurez-vous d'avoir :

- Création d'un Amazon Kendra index et, si vous utilisez l'API, notez l'ID de l'index.
- Vous avez créé un IAM rôle pour votre source de données et, si vous utilisez l'API, notez l'ARN du IAM rôle.

> **Note**
>
> Si vous modifiez votre type d'authentification et vos informations d'identification, vous devez mettre à jour votre IAM rôle pour accéder au bon identifiant AWS Secrets Manager secret.

- Stockez vos informations d'authentification Confluence dans un AWS Secrets Manager secret et, si vous utilisez l'API, notez l'ARN du secret.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

Si vous n'avez pas de IAM rôle ou de secret existant, vous pouvez utiliser la console pour créer un nouveau IAM rôle et un nouveau Secrets Manager secret lorsque vous connectez votre source de données Confluence à Amazon Kendra. Si vous utilisez l'API, vous devez fournir l'ARN d'un IAM rôle et d'un Secrets Manager secret existants, ainsi qu'un identifiant d'index.

Instructions de connexion

Pour vous connecter Amazon Kendra à votre source de données Confluence, vous devez fournir les détails de vos informations d'identification Confluence afin de Amazon Kendra pouvoir accéder à vos données. Si vous n'avez pas encore configuré Confluence pour Amazon Kendra voirPrérequis.

Console

Pour vous connecter Amazon Kendra à Confluence

1. Connectez-vous à la console AWS de gestion et Amazon Kendra ouvrez-la.
2. Dans le volet de navigation de gauche, choisissez Index, puis choisissez l'index que vous souhaitez utiliser dans la liste des index.

> **Note**
>
> Vous pouvez choisir de configurer ou de modifier vos paramètres de contrôle d'accès utilisateur dans les paramètres de l'index.

3. Sur la page de démarrage, choisissez Ajouter une source de données.
4. Sur la page Ajouter une source de données, choisissez Confluence connector V1.0, puis sélectionnez Ajouter une source de données.
5. Sur la page Spécifier les détails de la source de données, entrez les informations suivantes :
   a. Dans Nom et description, pour Nom de la source de données : entrez le nom de votre source de données. Vous pouvez inclure des tirets, mais pas des espaces.
   b. (Facultatif) Description : entrez une description facultative pour votre source de données.
   c. Dans la langue par défaut : choisissez une langue pour filtrer vos documents pour l'index. Sauf indication contraire, la langue par défaut est l'anglais. La langue spécifiée dans les métadonnées du document remplace la langue sélectionnée.
   d. Dans Balises, pour Ajouter une nouvelle balise : incluez des balises facultatives pour rechercher et filtrer vos ressources ou suivre vos AWS coûts.
   e. Choisissez Suivant.
6. Sur la page Définir l'accès et la sécurité, entrez les informations suivantes :
   a. Choisissez entre Confluence Cloud et Confluence Server.
   b. Si vous choisissez Confluence Cloud, saisissez les informations suivantes :

i. URL de confluence : votre URL de confluence.

ii. AWS Secrets Manager secret —Choisissez un secret existant ou créez-en un nouveau Secrets Manager pour stocker vos identifiants d'authentification Confluence. Si vous choisissez de créer un nouveau secret, une fenêtre AWS Secrets Manager secrète s'ouvre.

- Entrez les informations suivantes dans la fenêtre Créer un AWS Secrets Manager secret :

I. Nom secret : le nom de votre secret. Le préfixe « AmazonKendra - Confluence-» est automatiquement ajouté à votre nom secret.

II. Pour le nom d'utilisateur et le mot de passe : entrez votre nom d'utilisateur et votre mot de passe Confluence.

III. Choisissez Enregistrer l'authentification.

c. Si vous choisissez Confluence Server, entrez les informations suivantes :

i. URL Confluence : votre nom d'utilisateur et votre mot de passe Confluence.

ii. (Facultatif) Pour le proxy Web, entrez les informations suivantes :

A. Nom d'hôte : nom d'hôte de votre compte Confluence.

B. Numéro de port —Port utilisé par le protocole de transport d'URL de l'hôte.

iii. Pour l'authentification, choisissez l'authentification de base ou le jeton d'accès personnel (Confluence Server uniquement).

iv. AWS Secrets Manager secret —Choisissez un secret existant ou créez-en un nouveau Secrets Manager pour stocker vos identifiants d'authentification Confluence. Si vous choisissez de créer un nouveau secret, une fenêtre AWS Secrets Manager secrète s'ouvre.

- Entrez les informations suivantes dans la fenêtre Créer un AWS Secrets Manager secret :

I. Nom secret : le nom de votre secret. Le préfixe « AmazonKendra - Confluence-» est automatiquement ajouté à votre nom secret.

II. Pour le nom d'utilisateur et le mot de passe : entrez les valeurs d'identification d'authentification que vous avez configurées dans Confluence. Si vous utilisez l'authentification de base, utilisez votre nom

d'utilisateur Confluence (identifiant e-mail) et votre mot de passe (jeton API). Si vous utilisez un jeton d'accès personnel, saisissez les détails du jeton d'accès personnel que vous avez configuré dans le compte Confluence.

III. Enregistrez et ajoutez votre secret.

d. IAM rôle —Choisissez un IAM rôle existant ou créez-en un nouveau IAM pour accéder aux informations d'identification de votre référentiel et indexer le contenu.

> **Note**
>
> IAM les rôles utilisés pour les index ne peuvent pas être utilisés pour les sources de données. Si vous ne savez pas si un rôle existant est utilisé pour un index ou une FAQ, choisissez Créer un nouveau rôle pour éviter les erreurs.

e. Choisissez Suivant.

7. Sur la page Configurer les paramètres de synchronisation, entrez les informations suivantes :

   a. Pour Inclure les espaces personnels et Inclure les espaces archivés, choisissez les types d'espaces facultatifs à inclure dans cette source de données.

   b. Pour une configuration supplémentaire : spécifiez des modèles d'expressions régulières pour inclure ou exclure certains contenus. Vous pouvez ajouter jusqu'à 100 motifs.

   c. Vous pouvez également choisir d'explorer les pièces jointes dans les espaces choisis.

   d. Dans Synchroniser le calendrier d'exécution, pour Fréquence : choisissez la fréquence de synchronisation avec votre source de données. Amazon Kendra

   e. Choisissez Suivant.

8. Sur la page Définir les mappages de champs, entrez les informations suivantes :

   a. Pour Space, Page, Blog : sélectionnez parmi les champs de source de données par défaut Amazon Kendra générés ou parmi les mappages de champs suggérés supplémentaires pour ajouter des champs d'index.

   b. Ajouter un champ —Pour ajouter des champs de source de données personnalisés afin de créer un nom de champ d'index à mapper et le type de données du champ.

   c. Choisissez Suivant.

9. Sur la page Réviser et créer, vérifiez que les informations saisies sont correctes, puis sélectionnez Ajouter une source de données. Vous pouvez également choisir de modifier vos

informations depuis cette page. Votre source de données apparaîtra sur la page Sources de données une fois que la source de données aura été ajoutée avec succès.

API

Pour vous connecter Amazon Kendra à Confluence

Vous devez spécifier les éléments suivants à l'aide de l'ConfluenceConfigurationAPI :

- Version Confluence —Spécifiez la version de l'instance Confluence que vous utilisez en tant que ou. `CLOUD` `SERVER`
- Nom de ressource Amazon (ARN) secret : indiquez le nom de ressource Amazon (ARN) d'un Secrets Manager secret contenant vos informations d'authentification Confluence.

  Si vous utilisez Confluence Server, vous pouvez utiliser soit votre nom d'utilisateur et votre mot de passe Confluence, soit votre jeton d'accès personnel comme identifiants d'authentification.

  Si vous utilisez votre nom d'utilisateur et votre mot de passe Confluence comme informations d'authentification, vous stockez les informations d'identification suivantes sous forme de structure JSON dans votre Secrets Manager secret :

```
{
    "username": "user name",
    "password": "password"
}
```

  Si vous utilisez un jeton d'accès personnel pour vous connecter à Confluence Server Amazon Kendra, vous stockez les informations d'identification suivantes sous forme de structure JSON dans votre Secrets Manager secret :

```
{
    "patToken": "personal access token"
}
```

  Si vous utilisez Confluence Cloud, vous utilisez votre nom d'utilisateur Confluence et un jeton d'API, configuré dans Confluence, comme mot de passe. Vous stockez les informations d'identification suivantes sous forme de structure JSON dans votre Secrets Manager secret :

```
{
```

```
    "username": "user name",
    "password": "API token"
}
```

- IAM role —Spécifiez à quel `RoleArn` moment vous appelez `CreateDataSource` pour fournir à un IAM rôle les autorisations d'accéder à votre Secrets Manager secret et d'appeler le public requis APIs pour le connecteur Confluence et. Amazon Kendra Pour plus d'informations, consultez la section IAM Rôles des sources de données Confluence.

Vous pouvez également ajouter les fonctionnalités optionnelles suivantes :

- Proxy Web : s'il faut se connecter à votre instance d'URL Confluence via un proxy Web. Vous pouvez utiliser cette option pour Confluence Server.
- (Pour Confluence Server uniquement) Virtual Private Cloud (VPC) —Spécifiez dans le `VpcConfiguration` cadre de la configuration de la source de données. Consultez la section Configuration Amazon Kendra pour utiliser un VPC.
- Filtres d'inclusion et d'exclusion : spécifiez des modèles d'expressions régulières pour inclure ou exclure certains espaces, articles de blog, pages, espaces et pièces jointes. Si vous choisissez d'indexer les pièces jointes, seules les pièces jointes des pages et blogs indexés sont indexées.

> **Note**
>
> La plupart des sources de données utilisent des modèles d'expressions régulières, qui sont des modèles d'inclusion ou d'exclusion appelés filtres. Si vous spécifiez un filtre d'inclusion, seul le contenu correspondant au filtre d'inclusion est indexé. Tout document qui ne correspond pas au filtre d'inclusion n'est pas indexé. Si vous spécifiez un filtre d'inclusion et d'exclusion, les documents correspondant au filtre d'exclusion ne sont pas indexés, même s'ils correspondent au filtre d'inclusion.

- Mappages de champs : choisissez de mapper les champs de votre source de données Confluence à vos Amazon Kendra champs d'index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.

> **Note**
>
> Le champ du corps du document ou l'équivalent du corps du document pour vos documents est requis pour Amazon Kendra effectuer une recherche dans vos

documents. Vous devez associer le nom du champ du corps du document dans votre source de données au nom du champ d'index`_document_body`. Tous les autres champs sont facultatifs.

- Filtrage du contexte utilisateur et contrôle d'accèsAmazon Kendra : analyse la liste de contrôle d'accès (ACL) de vos documents, si vous disposez d'une ACL pour vos documents. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.

En savoir plus

Pour en savoir plus sur l'intégration Amazon Kendra à votre source de données Confluence, consultez :

- Configuration de votre Amazon Kendra connecteur Confluence Server

Connecteur de source de données personnalisé

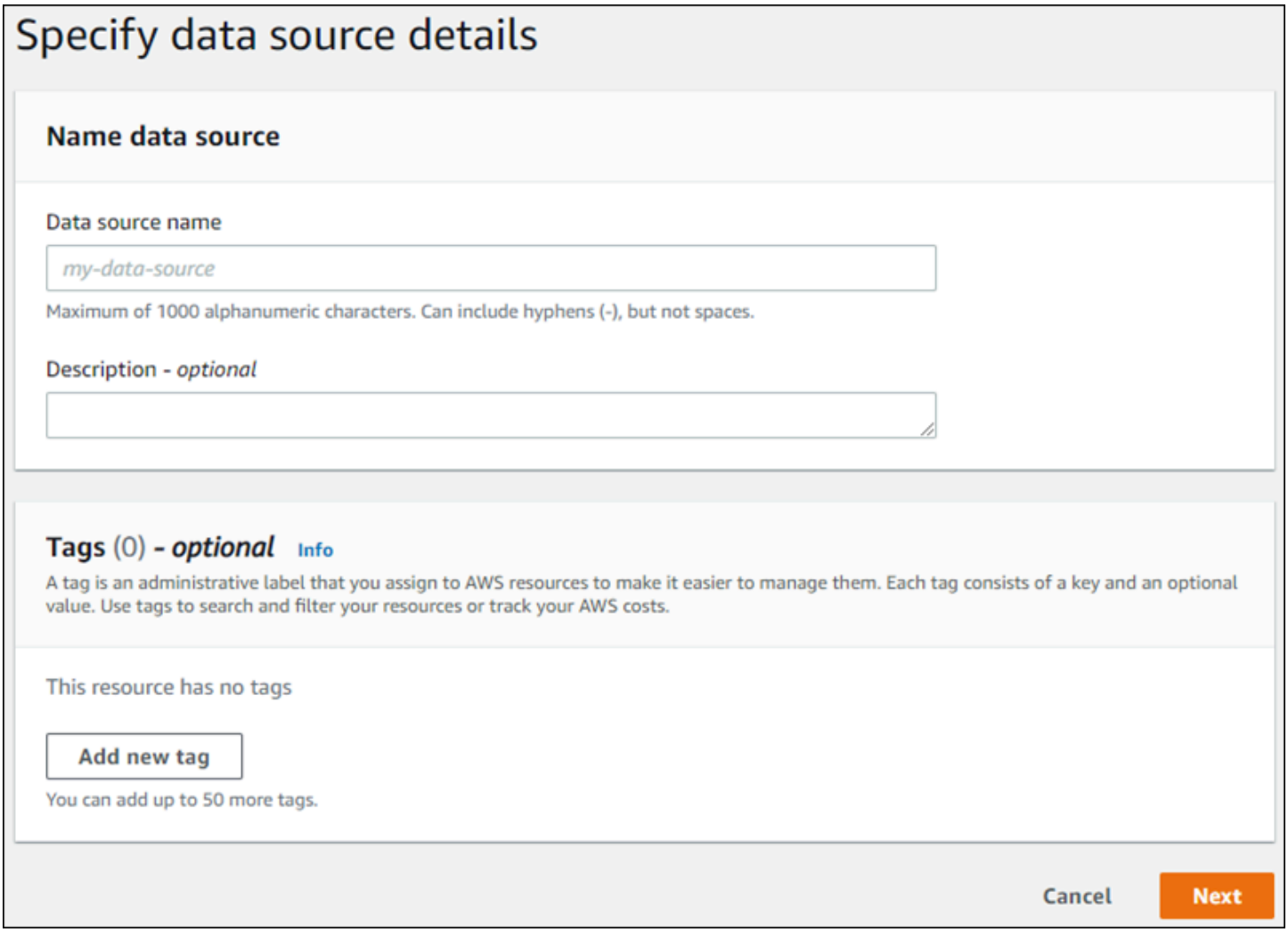
Utilisez une source de données personnalisée lorsque vous disposez d'un référentiel qui Amazon Kendra ne fournit pas encore de connecteur de source de données pour. Vous pouvez l'utiliser pour consulter les mêmes statistiques d'historique d'exécution que celles fournies par Amazon Kendra les sources de données, même lorsque vous ne pouvez pas utiliser les sources Amazon Kendra de données pour synchroniser vos référentiels. Utilisez-le pour créer une expérience de surveillance de synchronisation cohérente entre les sources de Amazon Kendra données et les sources personnalisées. Plus précisément, utilisez une source de données personnalisée pour voir les métriques de synchronisation d'un connecteur de source de données que vous avez créé à l'aide du BatchPutDocumentet BatchDeleteDocument APIs.

Pour résoudre les problèmes liés à votre connecteur de source de données personnalisé Amazon Kendra, consultez. Dépannage des sources de données

Lorsque vous créez une source de données personnalisée, vous contrôlez totalement la manière dont les documents à indexer sont sélectionnés. Amazon Kendra fournit uniquement des informations métriques que vous pouvez utiliser pour surveiller vos tâches de synchronisation de sources de données. Vous devez créer et exécuter le robot d'exploration qui détermine les documents indexés par votre source de données.

Vous devez spécifier le titre principal de vos documents à l'aide de l'objet Document, `_source_uri` DocumentAttributeafin de l'avoir `DocumentTitle` et de l'`DocumentURI`inclure dans la réponse du `Query` résultat.

Vous créez un identifiant pour votre source de données personnalisée à l'aide de la console ou de l'CreateDataSourceAPI. Pour utiliser la console, donnez un nom à votre source de données, éventuellement une description et des balises de ressources. Une fois la source de données créée, un identifiant de source de données est affiché. Copiez cet ID à utiliser lors de la synchronisation de la source de données avec l'index.



Vous pouvez également créer une source de données personnalisée à l'aide de l'`CreateDataSource`API. L'API renvoie un identifiant à utiliser lors de la synchronisation de la source de données. Lorsque vous utilisez l'`CreateDataSource`API pour créer une source de données personnalisée, vous ne pouvez pas définir `Configuration` les `Schedule`

paramètres `RoleArn` ou. Si vous définissez ces paramètres, Amazon Kendra renvoie une `ValidationException` exception.

Pour utiliser une source de données personnalisée, créez une application chargée de mettre à jour l' Amazon Kendra index. L'application dépend d'un robot d'exploration que vous créez. Le robot lit les documents de votre dépôt et détermine ceux à qui ils doivent être envoyés Amazon Kendra. Votre application doit suivre les étapes suivantes :

1. Explorez votre référentiel et dressez une liste des documents ajoutés, mis à jour ou supprimés.
2. Appelez l'StartDataSourceSyncJobAPI pour signaler qu'une tâche de synchronisation est en cours de démarrage. Vous fournissez un ID de source de données pour identifier la source de données en cours de synchronisation. Amazon Kendra renvoie un ID d'exécution pour identifier une tâche de synchronisation particulière.
3. Appelez l'BatchDeleteDocumentAPI pour supprimer des documents de l'index. Vous fournissez l'ID de source de données et l'ID d'exécution pour identifier la source de données synchronisée et la tâche à laquelle cette mise à jour est associée.
4. Appelez l'StopDataSourceSyncJobAPI pour signaler la fin de la tâche de synchronisation. Après avoir appelé l'`StopDataSourceSyncJob`API, l'ID d'exécution associé n'est plus valide.
5. Appelez l'ListDataSourceSyncJobsAPI avec l'index et les identifiants de source de données pour répertorier les tâches de synchronisation pour la source de données et pour voir les métriques relatives aux tâches de synchronisation.

Une fois que vous avez terminé une tâche de synchronisation, vous pouvez en démarrer une nouvelle. Il peut s'écouler un certain temps avant que tous les documents soumis soient ajoutés à l'index. Utilisez l'`ListDataSourceSyncJobs`API pour connaître l'état de la tâche de synchronisation. Si le résultat `Status` de la tâche de synchronisation est le cas`SYNCING_INDEXING`, certains documents sont toujours en cours d'indexation. Vous pouvez démarrer une nouvelle tâche de synchronisation lorsque le statut de la tâche précédente est `FAILED` ou`SUCCEEDED`.

Après avoir appelé l'`StopDataSourceSyncJob`API, vous ne pouvez pas utiliser d'identifiant de tâche de synchronisation dans un appel au `BatchPutDocument` ou `BatchDeleteDocument` APIs. Dans ce cas, tous les documents soumis sont renvoyés dans le message de `FailedDocuments` réponse de l'API.

Attributs requis

Lorsque vous soumettez un document à Amazon Kendra l'aide de l'`BatchPutDocument`API, chaque document nécessite deux attributs pour identifier la source de données et le cycle de synchronisation auxquels il appartient. Vous devez fournir les deux attributs suivants pour mapper correctement les documents de votre source de données personnalisée vers un Amazon Kendra index :

- `_data_source_id`: l'identifiant de la source de données. Cela est renvoyé lorsque vous créez la source de données à l'aide de la console ou de l'`CreateDataSource`API.
- `_data_source_sync_job_execution_id`: identifiant de la synchronisation exécutée. Cela est renvoyé lorsque vous lancez la synchronisation de l'index avec l'`StartDataSourceSyncJob`API.

Voici le JSON requis pour indexer un document à l'aide d'une source de données personnalisée.

```
{
    "Documents": [
        {
            "Attributes": [
                {
                    "Key": "_data_source_id",
                    "Value": {
                        "StringValue": "data source identifier"
                    }
                },
                {
                    "Key": "_data_source_sync_job_execution_id",
                    "Value": {
                        "StringValue": "sync job identifier"
                    }
                }
            ],
            "Blob": "document content",
            "ContentType": "content type",
            "Id": "document identifier",
            "Title": "document title"
        }
    ],
    "IndexId": "index identifier",
    "RoleArn": "IAM role ARN"
}
```

Lorsque vous supprimez un document de l'index à l'aide de l'`BatchDeleteDocument`API, vous devez spécifier les deux champs suivants dans le `DataSourceSyncJobMetricTarget` paramètre :

- `DataSourceId`: l'identifiant de la source de données. Cela est renvoyé lorsque vous créez la source de données à l'aide de la console ou de l'`CreateDataSource`API.
- `DataSourceSyncJobId`: identifiant de la synchronisation exécutée. Cela est renvoyé lorsque vous lancez la synchronisation de l'index avec l'`StartDataSourceSyncJob`API.

Voici le JSON requis pour supprimer un document de l'index à l'aide de l'`BatchDeleteDocument`API.

```
{
    "DataSourceSyncJobMetricTarget": {
        "DataSourceId": "data source identifier",
        "DataSourceSyncJobId": "sync job identifier"
    },
    "DocumentIdList": [
        "document identifier"
    ],
    "IndexId": "index identifier"
}
```

Affichage des métriques

Une fois qu'une tâche de synchronisation est terminée, vous pouvez utiliser l'DataSourceSyncJobMetricsAPI pour obtenir les métriques associées à la tâche de synchronisation. Utilisez-le pour surveiller les synchronisations de vos sources de données personnalisées.

Si vous soumettez le même document plusieurs fois, soit dans le cadre de l'`BatchPutDocument`API, soit dans le cadre de l'`BatchDeleteDocument`API, soit s'il est soumis à la fois pour ajout et suppression, le document n'est pris en compte qu'une seule fois dans les statistiques.

- `DocumentsAdded`: le nombre de documents soumis à l'aide de l'`BatchPutDocument`API associée à cette tâche de synchronisation ajoutés à l'index pour la première fois. Si un document est soumis pour ajout plusieurs fois lors d'une synchronisation, il n'est pris en compte qu'une seule fois dans les métriques.
- `DocumentsDeleted`: le nombre de documents soumis à l'aide de l'`BatchDeleteDocument`API associée à cette tâche de synchronisation supprimés de l'index. Si un document est soumis pour

suppression plusieurs fois lors d'une synchronisation, il n'est pris en compte qu'une seule fois dans les statistiques.

- `DocumentsFailed`: le nombre de documents associés à cette tâche de synchronisation dont l'indexation a échoué. Il s'agit de documents qui ont été acceptés Amazon Kendra pour indexation mais qui n'ont pas pu être indexés ou supprimés. Si un document n'est pas accepté par Amazon Kendra, son identifiant est renvoyé dans la propriété de `FailedDocuments` réponse du `BatchPutDocument` et `BatchDeleteDocument` APIs.
- `DocumentsModified`: le nombre de documents modifiés soumis à l'aide de l'`BatchPutDocument`API associée à cette tâche de synchronisation qui ont été modifiés dans l' Amazon Kendra index.

Amazon Kendra émet également des Amazon CloudWatch métriques lors de l'indexation des documents. Pour plus d'informations, consultez la section Surveillance Amazon Kendra avec Amazon CloudWatch.

Amazon Kendra ne renvoie pas la `DocumentsScanned` métrique pour les sources de données personnalisées. Il émet également les CloudWatch métriques répertoriées dans le document Métriques pour les sources de Amazon Kendra données.

En savoir plus

Pour en savoir plus sur l'intégration Amazon Kendra à votre source de données personnalisée, consultez :

- Ajouter des sources de données personnalisées à Amazon Kendra

Source de données personnalisée (Java)

Le code suivant fournit un exemple d'implémentation d'une source de données personnalisée à l'aide de Java. Le programme crée d'abord une source de données personnalisée, puis synchronise les documents récemment ajoutés à l'index avec la source de données personnalisée.

Le code suivant illustre la création et l'utilisation d'une source de données personnalisée. Lorsque vous utilisez une source de données personnalisée dans votre application, vous n'avez pas besoin de créer une nouvelle source de données (processus ponctuel) chaque fois que vous synchronisez votre index avec votre source de données. Vous utilisez l'ID d'index et l'ID de source de données pour synchroniser vos données.

```
package com.amazonaws.kendra;

import java.util.concurrent.TimeUnit;
import software.amazon.awssdk.services.kendra.KendraClient;
import csoftware.amazon.awssdk.services.kendra.model.BatchPutDocumentRequest;
import csoftware.amazon.awssdk.services.kendra.model.BatchPutDocumentResponse;
import software.amazon.awssdk.services.kendra.model.CreateDataSourceRequest;
import software.amazon.awssdk.services.kendra.model.CreateDataSourceResponse;
import software.amazon.awssdk.services.kendra.model.DataSourceType;
import software.amazon.awssdk.services.kendra.model.Document;
import software.amazon.awssdk.services.kendra.model.ListDataSourceSyncJobsRequest;
import software.amazon.awssdk.services.kendra.model.ListDataSourceSyncJobsResponse;
import software.amazon.awssdk.services.kendra.model.StartDataSourceSyncJobRequest;
import software.amazon.awssdk.services.kendra.model.StartDataSourceSyncJobResponse;
import software.amazon.awssdk.services.kendra.model.StopDataSourceSyncJobRequest;
import software.amazon.awssdk.services.kendra.model.StopDataSourceSyncJobResponse;

public class SampleSyncForCustomDataSource {
  public static void main(String[] args) {
    KendraClient kendra = KendraClient.builder().build();

    String myIndexId = "yourIndexId";
    String dataSourceName = "custom data source";
    String dataSourceDescription = "Amazon Kendra custom data source connector"

    // Create custom data source
    CreateDataSourceRequest createDataSourceRequest = CreateDataSourceRequest
        .builder()
        .indexId(myIndexId)
        .name(dataSourceName)
        .description(dataSourceDescription)
        .type(DataSourceType.CUSTOM)
        .build();

    CreateDataSourceResponse createDataSourceResponse =
 kendra.createDataSource(createDataSourceRequest);
    System.out.println(String.format("Response of creating data source: %s",
 createDataSourceResponse));

    // Get the data source ID from createDataSourceResponse
    String dataSourceId = createDataSourceResponse.Id();

    // Wait for the custom data source to become active
```

```
   System.out.println(String.format("Waiting for Amazon Kendra to create the data
source %s", dataSourceId));
   // You can use the DescribeDataSource API to check the status
   DescribeDataSourceRequest describeDataSourceRequest = DescribeDataSourceRequest
        .builder()
        .indexId(myIndexId)
        .id(dataSourceId)
        .build();

   while (true) {
        DescribeDataSourceResponse describeDataSourceResponse =
kendra.describeDataSource(describeDataSourceRequest);

        DataSourceStatus status = describeDataSourceResponse.status();
        System.out.println(String.format("Creating data source. Status: %s", status));
        if (status != DataSourceStatus.CREATING) {
            break;
        }

        TimeUnit.SECONDS.sleep(60);
   }

   // Start syncing yor data source by calling StartDataSourceSyncJob and providing
your index ID
   // and your custom data source ID
   System.out.println(String.format("Synchronize the data source %s", dataSourceId));
   StartDataSourceSyncJobRequest startDataSourceSyncJobRequest =
StartDataSourceSyncJobRequest
        .builder()
        .indexId(myIndexId)
        .id(dataSourceId)
        .build();
   StartDataSourceSyncJobResponse startDataSourceSyncJobResponse =
kendra.startDataSourceSyncJob(startDataSourceSyncJobRequest);

   // Get the  sync job execution ID from startDataSourceSyncJobResponse
   String executionId = startDataSourceSyncJobResponse.ExecutionId();
System.out.println(String.format("Waiting for the data source to sync with the index
%s for execution ID %s", indexId, startDataSourceSyncJobResponse.executionId()));

   // Add 2 documents uploaded to S3 bucket to your index using the BatchPutDocument
API
   // The added documents should sync with your custom data source
   Document pollyDoc = Document
```

```
        .builder()
        .s3Path(
            S3Path.builder()
            .bucket("amzn-s3-demo-bucket")
            .key("what_is_Amazon_Polly.docx")
            .build())
        .title("What is Amazon Polly?")
        .id("polly_doc_1")
        .build();

    Document rekognitionDoc = Document
        .builder()
        .s3Path(
            S3Path.builder()
            .bucket("amzn-s3-demo-bucket")
            .key("what_is_amazon_rekognition.docx")
            .build())
        .title("What is Amazon rekognition?")
        .id("rekognition_doc_1")
        .build();

    BatchPutDocumentRequest batchPutDocumentRequest = BatchPutDocumentRequest
        .builder()
        .indexId(myIndexId)
        .documents(pollyDoc, rekognitionDoc)
        .build();

    BatchPutDocumentResponse result = kendra.batchPutDocument(batchPutDocumentRequest);
    System.out.println(String.format("BatchPutDocument result: %s", result));

    // Once custom data source synced, stop the sync job using the
StopDataSourceSyncJob API
    StopDataSourceSyncJobResponse stopDataSourceSyncJobResponse =
kendra.stopDataSourceSyncJob(
        StopDataSourceSyncJobRequest()
            .indexId(myIndexId)
            .id(dataSourceId)
    );

// List your sync jobs
    ListDataSourceSyncJobsRequest listDataSourceSyncJobsRequest =
ListDataSourceSyncJobsRequest
        .builder()
        .indexId(myIndexId)
```

```
        .id(dataSourceId)
        .build();

    while (true) {
        ListDataSourceSyncJobsResponse listDataSourceSyncJobsResponse =
 kendra.listDataSourceSyncJobs(listDataSourceSyncJobsRequest);
        DataSourceSyncJob job = listDataSourceSyncJobsResponse.history().get(0);
        System.out.println(String.format("Status: %s", job.status()));
    }
  }
}
```

Dropbox

Dropbox est un service d'hébergement de fichiers qui propose des services de stockage dans le cloud, d'organisation de documents et de création de modèles de documents. Si vous êtes un utilisateur de Dropbox, vous pouvez l'utiliser Amazon Kendra pour indexer vos fichiers Dropbox, Dropbox Paper, les modèles Dropbox Paper et les raccourcis enregistrés vers des pages Web. Vous pouvez également configurer Amazon Kendra pour indexer des fichiers Dropbox spécifiques, Dropbox Paper, des modèles Dropbox Paper et des raccourcis enregistrés vers des pages Web.

Amazon Kendra prend en charge Dropbox et Dropbox Advanced pour Dropbox Business.

Vous pouvez vous connecter Amazon Kendra à votre source de données Dropbox à l'aide de la Amazon Kendra console et de l'TemplateConfigurationAPI.

Pour résoudre les problèmes liés à votre connecteur de source de données Amazon Kendra Dropbox, consultezDépannage des sources de données.

Rubriques

- Fonctionnalités prises en charge
- Prérequis
- Instructions de connexion
- En savoir plus
- Remarques

Fonctionnalités prises en charge

Amazon Kendra Le connecteur de source de données Dropbox prend en charge les fonctionnalités suivantes :

- Mappages de champs
- Contrôle d'accès utilisateur
- Filtres d'inclusion/exclusion
- Synchronisation complète et incrémentielle du contenu
- Cloud privé virtuel (VPC)

Prérequis

Avant de pouvoir Amazon Kendra indexer votre source de données Dropbox, apportez ces modifications à votre Dropbox et à vos AWS comptes.

Dans Dropbox, assurez-vous que vous disposez des éléments suivants :

- Vous avez créé un compte Dropbox Advanced et configuré un utilisateur administrateur.
- Configuration d'une application Dropbox avec un nom d'application unique, activation de Scoped Access. Consultez la documentation Dropbox sur la création d'une application.
- Activation des autorisations Dropbox complètes sur la console Dropbox et ajout des autorisations suivantes :
  - fichiers.contenu.read
  - files.metadata.read
  - partager.lire
  - fichier_requests.read
  - groups.read
  - team_info.read
  - team_data.content.read
- Vous avez pris note de votre clé d'application Dropbox, du secret de l'application Dropbox et du jeton d'accès Dropbox pour les informations d'authentification de base.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

- Vous avez configuré et copié un jeton d'accès Oauth 2.0 temporaire pour votre application Dropbox. Ce jeton est temporaire et expire au bout de 4 heures. Consultez la documentation Dropbox sur OAuth l'authentification.

> **Note**
>
> Il est recommandé de créer un jeton d'accès d'actualisation Dropbox qui n'expire jamais, plutôt que de vous fier à un jeton d'accès unique qui expire au bout de 4 heures. Un jeton d'accès à l'actualisation est permanent et n'expire jamais afin que vous puissiez continuer à synchroniser votre source de données à l'avenir.

- Recommandé : Configurez un jeton d'actualisation permanent Dropbox qui n'expire jamais Amazon Kendra pour vous permettre de continuer à synchroniser votre source de données sans aucune interruption. Consultez la documentation Dropbox sur les jetons d'actualisation.
- Il est vérifié que chaque document est unique dans Dropbox et dans les autres sources de données que vous prévoyez d'utiliser pour le même index. Chaque source de données que vous souhaitez utiliser pour un index ne doit pas contenir le même document dans toutes les sources de données. IDs Les documents sont globaux par rapport à un index et doivent être uniques par index.

Dans votre Compte AWS, assurez-vous d'avoir :

- Création d'un Amazon Kendra index et, si vous utilisez l'API, notez l'ID de l'index.
- Vous avez créé un IAM rôle pour votre source de données et, si vous utilisez l'API, notez l'ARN du IAM rôle.

> **Note**
>
> Si vous modifiez votre type d'authentification et vos informations d'identification, vous devez mettre à jour votre IAM rôle pour accéder au bon identifiant AWS Secrets Manager secret.

- Stockez vos informations d'authentification Dropbox dans un AWS Secrets Manager secret et, si vous utilisez l'API, notez l'ARN du secret.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

Si vous n'avez aucun IAM rôle ou secret existant, vous pouvez utiliser la console pour créer un nouveau IAM rôle et un nouveau Secrets Manager secret lorsque vous connectez votre source de données Dropbox à Amazon Kendra. Si vous utilisez l'API, vous devez fournir l'ARN d'un IAM rôle et d'un Secrets Manager secret existants, ainsi qu'un identifiant d'index.

Instructions de connexion

Pour vous connecter Amazon Kendra à votre source de données Dropbox, vous devez fournir les informations nécessaires sur celle-ci afin de Amazon Kendra pouvoir accéder à vos données. Si vous n'avez pas encore configuré Dropbox pour Amazon Kendra, consultezPrérequis.

Console

Pour vous connecter Amazon Kendra à Dropbox

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la.
2. Dans le volet de navigation de gauche, choisissez Index, puis choisissez l'index que vous souhaitez utiliser dans la liste des index.

> **Note**
>
> Vous pouvez choisir de configurer ou de modifier vos paramètres de contrôle d'accès utilisateur dans les paramètres de l'index.

3. Sur la page de démarrage, choisissez Ajouter une source de données.
4. Sur la page Ajouter une source de données, choisissez le connecteur Dropbox, puis sélectionnez Ajouter un connecteur. Si vous utilisez la version 2 (le cas échéant), choisissez le connecteur Dropbox avec le tag « V2.0 ».
5. Sur la page Spécifier les détails de la source de données, entrez les informations suivantes :
   a. Dans Nom et description, pour Nom de la source de données : entrez le nom de votre source de données. Vous pouvez inclure des tirets, mais pas des espaces.
   b. (Facultatif) Description : entrez une description facultative pour votre source de données.
   c. Dans la langue par défaut : choisissez une langue pour filtrer vos documents pour l'index. Sauf indication contraire, la langue par défaut est l'anglais. La langue spécifiée dans les métadonnées du document remplace la langue sélectionnée.
   d. Dans Balises, pour Ajouter une nouvelle balise : incluez des balises facultatives pour rechercher et filtrer vos ressources ou suivre vos AWS coûts.
   e. Choisissez Suivant.
6. Sur la page Définir l'accès et la sécurité, entrez les informations suivantes :
   a. Autorisation : activez ou désactivez les informations de la liste de contrôle d'accès (ACL) pour vos documents, si vous disposez d'une ACL et que vous souhaitez l'utiliser pour le contrôle d'accès. L'ACL indique les documents auxquels les utilisateurs et les groupes peuvent accéder. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.
   b. Type de jeton d'authentification : choisissez un jeton permanent (recommandé) ou un jeton d'accès temporaire.
   c. AWS Secrets Manager secret —Choisissez un secret existant ou créez-en un nouveau Secrets Manager pour stocker vos informations d'authentification Dropbox. Si vous choisissez de créer un nouveau secret, une fenêtre AWS Secrets Manager secrète s'ouvre.

i. Entrez les informations suivantes dans la fenêtre Créer un AWS Secrets Manager secret :

A. Nom secret : le nom de votre secret. Le préfixe « AmazonKendra -Dropbox-» est automatiquement ajouté à votre nom secret.

B. Pour les informations relatives à la clé d'application, au secret de l'application et au jeton (permanent ou temporaire) : entrez les valeurs d'identification d'authentification configurées dans Dropbox.

ii. Enregistrez et ajoutez votre secret.

d. Virtual Private Cloud (VPC) —Vous pouvez choisir d'utiliser un VPC. Dans ce cas, vous devez ajouter des sous-réseaux et des groupes de sécurité VPC.

e. Identity Crawler : spécifiez s'il faut activer l'explorateur Amazon Kendra d'identité. Le robot d'exploration d'identité utilise les informations de la liste de contrôle d'accès (ACL) de vos documents pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Si vous disposez d'une ACL pour vos documents et que vous choisissez de l'utiliser, vous pouvez également choisir d'activer le robot d'exploration Amazon Kendra d'identité pour configurer le filtrage des résultats de recherche par contexte utilisateur. Sinon, si le robot d'identification est désactivé, tous les documents peuvent être consultés publiquement. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'exploration d'identité est désactivé, vous pouvez également utiliser l'PutPrincipalMappingAPI pour télécharger les informations d'accès des utilisateurs et des groupes afin de filtrer le contexte utilisateur.

f. IAM rôle —Choisissez un IAM rôle existant ou créez-en un nouveau IAM pour accéder aux informations d'identification de votre référentiel et indexer le contenu.

> **Note**
>
> IAM les rôles utilisés pour les index ne peuvent pas être utilisés pour les sources de données. Si vous ne savez pas si un rôle existant est utilisé pour un index ou une FAQ, choisissez Créer un nouveau rôle pour éviter les erreurs.

g. Choisissez Suivant.

7. Sur la page Configurer les paramètres de synchronisation, entrez les informations suivantes :

a. Pour Sélectionner des entités ou des types de contenu : choisissez les entités Dropbox ou les types de contenu que vous souhaitez analyser.

b. Dans Configuration supplémentaire pour les modèles Regex : ajoutez des modèles d'expressions régulières pour inclure ou exclure certains fichiers.

c. Mode de synchronisation : choisissez la manière dont vous souhaitez mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation.

   - Synchronisation complète : Indexez tout le contenu fraîchement, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
   - Nouvelle synchronisation modifiée : Indexez uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
   - Synchronisation nouvelle, modifiée ou supprimée : Indexez uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

d. Dans Calendrier d'exécution de la synchronisation, pour Fréquence : choisissez la fréquence à laquelle vous souhaitez synchroniser le contenu de votre source de données et mettre à jour votre index.

e. Choisissez Suivant.

8. Sur la page Définir les mappages de champs, entrez les informations suivantes :

   a. Fichiers, Dropbox Paper et modèles Dropbox Paper : sélectionnez parmi les champs de source de données par défaut Amazon Kendra générés que vous souhaitez associer à votre index.

   b. Ajouter un champ —Pour ajouter des champs de source de données personnalisés afin de créer un nom de champ d'index à mapper et le type de données du champ.

   c. Choisissez Suivant.

9. Sur la page Réviser et créer, vérifiez que les informations saisies sont correctes, puis sélectionnez Ajouter une source de données. Vous pouvez également choisir de modifier vos informations depuis cette page. Votre source de données apparaîtra sur la page Sources de données une fois que la source de données aura été ajoutée avec succès.

API

Pour vous connecter Amazon Kendra à Dropbox

Vous devez spécifier un JSON du schéma de source de données à l'aide de l'TemplateConfigurationAPI. Vous devez fournir les informations suivantes :

- Source de données —Spécifiez le type de source de données tel que `DROPBOX` lorsque vous utilisez le schéma TemplateConfigurationJSON. Spécifiez également la source de données `TEMPLATE` lorsque vous appelez l'CreateDataSourceAPI.
- Mode de synchronisation : spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation. Vous pouvez choisir entre :
  - `FORCED_FULL_CRAWL`pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
  - `FULL_CRAWL`pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
  - `CHANGE_LOG`pour indexer uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
- Type de jeton d'accès : spécifiez si vous souhaitez utiliser un jeton d'accès permanent ou temporaire pour votre AWS Secrets Manager secret qui stocke vos informations d'authentification.

> **Note**
>
> Il est recommandé de créer un jeton d'accès d'actualisation qui n'expire jamais dans Dropbox plutôt que de vous fier à un jeton d'accès unique expirant au bout de 4 heures. Vous créez une application et un jeton d'accès actualisé dans la console de développement Dropbox, puis vous fournissez le jeton d'accès dans votre code secret.

- Nom de ressource Amazon (ARN) secret : indiquez le nom de ressource Amazon (ARN) d'un Secrets Manager secret contenant les informations d'authentification de votre compte Dropbox. Le secret est stocké dans une structure JSON avec les clés suivantes :

```
{
    "appKey": "Dropbox app key",
    "appSecret": "Dropbox app secret",
    "accesstoken": "temporary access token or refresh access token"
}
```

- Identity Crawler : spécifiez s'il faut activer l'explorateur Amazon Kendra d'identité. Le robot d'exploration d'identité utilise les informations de la liste de contrôle d'accès (ACL) de vos documents pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Si vous disposez d'une ACL pour vos documents et que vous choisissez de l'utiliser, vous pouvez également choisir d'activer le robot d'exploration Amazon Kendra d'identité pour configurer le filtrage des résultats de recherche par contexte utilisateur. Sinon, si le robot d'identification est désactivé, tous les documents peuvent être consultés publiquement. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'exploration d'identité est désactivé, vous pouvez également utiliser l'PutPrincipalMappingAPI pour télécharger les informations d'accès des utilisateurs et des groupes afin de filtrer le contexte utilisateur.
- IAM rôle : spécifiez à quel `RoleArn` moment vous appelez `CreateDataSource` pour accorder à un IAM rôle les autorisations d'accéder à votre Secrets Manager secret et pour appeler le public requis APIs pour le connecteur Dropbox et Amazon Kendra. Pour plus d'informations, consultez la section IAM Rôles des sources de données Dropbox.

Vous pouvez également ajouter les fonctionnalités optionnelles suivantes :

- Virtual Private Cloud (VPC) **`VpcConfiguration`** —Spécifiez le moment de votre appel. `CreateDataSource` Pour de plus amples informations, veuillez consulter Configuration Amazon Kendra pour utiliser un Amazon VPC.
- Types de documents/contenus : indiquez si vous souhaitez explorer les fichiers de votre Dropbox, les documents Dropbox Paper, les modèles Dropbox Paper et les raccourcis de pages Web stockés dans votre Dropbox.
- Filtres d'inclusion et d'exclusion : spécifiez si vous souhaitez inclure ou exclure certains fichiers.

> **Note**
>
> La plupart des sources de données utilisent des modèles d'expressions régulières, qui sont des modèles d'inclusion ou d'exclusion appelés filtres. Si vous spécifiez un filtre d'inclusion, seul le contenu correspondant au filtre d'inclusion est indexé. Tout document qui ne correspond pas au filtre d'inclusion n'est pas indexé. Si vous spécifiez un filtre d'inclusion et d'exclusion, les documents correspondant au filtre d'exclusion ne sont pas indexés, même s'ils correspondent au filtre d'inclusion.

- Liste de contrôle d'accès (ACL) : indiquez si vous souhaitez analyser les informations ACL de vos documents, si vous disposez d'une ACL et souhaitez l'utiliser pour le contrôle d'accès. L'ACL indique les documents auxquels les utilisateurs et les groupes peuvent accéder. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.
- Mappages de champs : choisissez de mapper les champs de votre source de données Dropbox à vos champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.

> **Note**
>
> Le champ du corps du document ou l'équivalent du corps du document pour vos documents est requis pour Amazon Kendra effectuer une recherche dans vos documents. Vous devez associer le nom du champ du corps du document dans votre source de données au nom du champ d'index`_document_body`. Tous les autres champs sont facultatifs.

Pour obtenir la liste des autres clés JSON importantes à configurer, consultez le schéma du modèle Dropbox.

En savoir plus

Pour en savoir plus sur l'intégration d'Amazon Kendra à votre source de données Dropbox, consultez :

- Indexez votre contenu Dropbox à l'aide du connecteur Dropbox pour Amazon Kendra

Remarques

- Lorsque les listes de contrôle d'accès (ACLs) sont activées, l'option « Synchroniser uniquement le contenu nouveau ou modifié » n'est pas disponible en raison des limites de l'API Dropbox. Nous vous recommandons d'utiliser plutôt les modes « Synchronisation complète » ou « Synchronisation du contenu nouveau, modifié ou supprimé », ou de les désactiver ACLs si vous devez utiliser ce mode de synchronisation.

Drupal

> **Note**
>
> Le connecteur Drupal reste entièrement pris en charge pour les clients existants jusqu'au 31 mai 2026. Bien que ce connecteur ne soit plus disponible pour les nouveaux utilisateurs, les utilisateurs actuels peuvent continuer à l'utiliser sans interruption. Nous faisons constamment évoluer notre portefeuille de connecteurs afin de proposer des solutions plus évolutives et personnalisables. Pour les intégrations futures, nous vous recommandons d'explorer le framework de connecteurs personnalisés Amazon Kendra [1], conçu pour prendre en charge un plus large éventail de cas d'utilisation en entreprise avec une flexibilité accrue.

Drupal est un système de gestion de contenu (CMS) open source que vous pouvez utiliser pour créer des sites Web et des applications Web. Vous pouvez l'utiliser Amazon Kendra pour indexer les éléments suivants dans Drupal :

- Contenu : articles, pages de base, blocs de base, types de contenu définis par l'utilisateur, types de blocs définis par l'utilisateur, types de contenu personnalisés, types de blocs personnalisés

- Commentaire : pour tous les types de contenu et de bloc
- Pièces jointes : pour tous les types de contenu et de blocs

Vous pouvez vous connecter Amazon Kendra à votre source de données Drupal à l'aide de la Amazon Kendra console ou de l'TemplateConfigurationAPI.

Pour résoudre les problèmes liés à votre connecteur de source de données Amazon Kendra Drupal, consultez. Dépannage des sources de données

Rubriques

- Fonctionnalités prises en charge
- Prérequis
- Instructions de connexion
- Remarques

Fonctionnalités prises en charge

Amazon Kendra Le connecteur de source de données Drupal prend en charge les fonctionnalités suivantes :

- Mappages de champs
- Filtrage du contexte utilisateur
- Filtres d'inclusion/exclusion
- Synchronisation complète et incrémentielle du contenu
- Cloud privé virtuel (VPC)

Prérequis

Avant de pouvoir Amazon Kendra indexer votre source de données Drupal, apportez ces modifications à votre Drupal et à vos comptes. AWS

Dans Drupal, assurez-vous d'avoir :

- Création d'un compte Drupal (Standard) Suite et d'un utilisateur avec un rôle d'administrateur.
- Vous avez copié le nom de votre site Drupal et configuré une URL d'hôte. Par exemple, *https://<hostname>/<drupalsitename>*.

- Identifiants d'authentification de base configurés contenant un nom d'utilisateur (nom d'utilisateur de connexion au site Web Drupal) et un mot de passe (mot de passe du site Web Drupal).
- Recommandé : configuration d'un jeton d'identification OAuth 2.0. Utilisez ce jeton avec votre mot de passe Drupal, votre identifiant client, votre secret client, votre nom d'utilisateur (nom d'utilisateur de connexion au site Web Drupal) et votre mot de passe (mot de passe du site Web Drupal) pour vous connecter à. Amazon Kendra
- Vous avez ajouté les autorisations suivantes à votre compte Drupal à l'aide d'un rôle d'administrateur :
  - administrer des blocs
  - administrer l'affichage de block_content
  - administrer les champs block_content
  - administrer l'affichage du formulaire block_content
  - administrer les vues
  - afficher les adresses e-mail des utilisateurs
  - voir son propre contenu non publié
  - voir les révisions de page
  - voir les révisions des articles
  - voir toutes les révisions
  - voir le thème de l'administration
  - accéder au contenu
  - aperçu du contenu d'accès
  - accéder aux commentaires
  - contenu de recherche
  - aperçu des fichiers d'accès
  - accéder aux liens contextuels

> **Note**
>
> S'il existe des types de contenu définis par l'utilisateur ou des types de blocs définis par l'utilisateur, ou si des vues et des blocs sont ajoutés au site Web Drupal, ils doivent disposer d'un accès administrateur.

Dans votre Compte AWS, assurez-vous d'avoir :

- Création d'un Amazon Kendra index et, si vous utilisez l'API, notez l'ID de l'index.
- Vous avez créé un IAM rôle pour votre source de données et, si vous utilisez l'API, notez l'ARN du IAM rôle.

> **Note**
>
> Si vous modifiez votre type d'authentification et vos informations d'identification, vous devez mettre à jour votre IAM rôle pour accéder au bon identifiant AWS Secrets Manager secret.

- Stockez vos informations d'authentification Drupal dans un AWS Secrets Manager secret et, si vous utilisez l'API, notez l'ARN du secret.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

Si vous n'avez pas de IAM rôle ou de secret existant, vous pouvez utiliser la console pour créer un nouveau IAM rôle et un nouveau Secrets Manager secret lorsque vous connectez votre source de données Drupal à. Amazon Kendra Si vous utilisez l'API, vous devez fournir l'ARN d'un IAM rôle et d'un Secrets Manager secret existants, ainsi qu'un identifiant d'index.

Instructions de connexion

Pour vous connecter Amazon Kendra à votre source de données Drupal, vous devez fournir les détails de vos informations d'identification Drupal afin de Amazon Kendra pouvoir accéder à vos données. Si vous n'avez pas encore configuré Drupal pour Amazon Kendra voir. Prérequis

Console

Pour vous connecter Amazon Kendra à Drupal

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la.
2. Dans le volet de navigation de gauche, choisissez Index, puis choisissez l'index que vous souhaitez utiliser dans la liste des index.

> **Note**
>
> Vous pouvez choisir de configurer ou de modifier vos paramètres de contrôle d'accès utilisateur dans les paramètres de l'index.

3. Sur la page de démarrage, choisissez Ajouter une source de données.
4. Sur la page Ajouter une source de données, choisissez le connecteur Drupal, puis sélectionnez Ajouter un connecteur. Si vous utilisez la version 2 (le cas échéant), choisissez le connecteur Drupal avec le tag « V2.0 ».
5. Sur la page Spécifier les détails de la source de données, entrez les informations suivantes :
   a. Dans Nom et description, pour Nom de la source de données : entrez le nom de votre source de données. Vous pouvez inclure des tirets, mais pas des espaces.
   b. (Facultatif) Description : entrez une description facultative pour votre source de données.
   c. Dans la langue par défaut : choisissez une langue pour filtrer vos documents pour l'index. Sauf indication contraire, la langue par défaut est l'anglais. La langue spécifiée dans les métadonnées du document remplace la langue sélectionnée.
   d. Dans Balises, pour Ajouter une nouvelle balise : incluez des balises facultatives pour rechercher et filtrer vos ressources ou suivre vos AWS coûts.
   e. Choisissez Suivant.
6. Sur la page Définir l'accès et la sécurité, entrez les informations suivantes :
   a. Dans Source, pour URL hôte : URL hôte de votre site Drupal. Par exemple, *https://<hostname>/<drupalsitename>*.
   b. Pour l'emplacement du certificat SSL : entrez le chemin d'accès au certificat SSL stocké dans votre Amazon S3 compartiment.
   c. Autorisation : activez ou désactivez les informations de la liste de contrôle d'accès (ACL) pour vos documents, si vous disposez d'une ACL et que vous souhaitez l'utiliser pour le contrôle d'accès. L'ACL indique les documents auxquels les utilisateurs et les groupes peuvent accéder. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.

d. Pour l'authentification : choisissez entre l'authentification de base et l'authentification OAuth 2.0 en fonction de votre cas d'utilisation.

e. AWS Secrets Manager secret —Choisissez un secret existant ou créez-en un nouveau Secrets Manager pour stocker vos informations d'authentification Drupal. Si vous choisissez de créer un nouveau secret, une fenêtre AWS Secrets Manager secrète s'ouvre.

   i. Entrez les informations suivantes dans la fenêtre Créer un AWS Secrets Manager secret :

      A. Si vous avez choisi l'authentification de base, entrez le nom secret, le nom d'utilisateur (nom d'utilisateur du site Drupal) et le mot de passe (mot de passe du site Drupal) que vous avez copiés, puis choisissez Enregistrer et ajouter un secret.

      B. Si vous avez choisi l'authentification OAuth 2.0, entrez un nom secret, un nom d'utilisateur (nom d'utilisateur du site Drupal), un mot de passe (mot de passe du site Drupal), un identifiant client et un secret client générés dans votre compte Drupal, puis choisissez Enregistrer et ajouter un secret.

   ii. Choisissez Enregistrer.

f. Virtual Private Cloud (VPC) —Vous pouvez choisir d'utiliser un VPC. Dans ce cas, vous devez ajouter des sous-réseaux et des groupes de sécurité VPC.

g. Identity Crawler : spécifiez s'il faut activer l'explorateur Amazon Kendra d'identité. Le robot d'exploration d'identité utilise les informations de la liste de contrôle d'accès (ACL) de vos documents pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Si vous disposez d'une ACL pour vos documents et que vous choisissez de l'utiliser, vous pouvez également choisir d'activer le robot d'exploration Amazon Kendra d'identité pour configurer le filtrage contextuel utilisateur des résultats de recherche. Sinon, si le robot d'identification est désactivé, tous les documents peuvent être consultés publiquement. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'exploration d'identité est désactivé, vous pouvez également utiliser l'PutPrincipalMappingAPI pour télécharger les informations d'accès des utilisateurs et des groupes afin de filtrer le contexte utilisateur.

h. IAM rôle —Choisissez un IAM rôle existant ou créez-en un nouveau IAM pour accéder aux informations d'identification de votre référentiel et indexer le contenu.

> **Note**
>
> IAM les rôles utilisés pour les index ne peuvent pas être utilisés pour les sources de données. Si vous ne savez pas si un rôle existant est utilisé pour un index ou une FAQ, choisissez Créer un nouveau rôle pour éviter les erreurs.

i. Choisissez Suivant.

7. Sur la page Configurer les paramètres de synchronisation, entrez les informations suivantes :

   a. Pour l'étendue de la synchronisation, choisissez l'une des options suivantes :

   > **Note**
   >
   > Lorsque vous choisissez d'explorer les articles, les pages de base et les blocs de base, leurs champs par défaut sont automatiquement synchronisés. Vous pouvez également choisir de synchroniser leurs commentaires, pièces jointes, champs personnalisés et autres entités personnalisées.

   - Pour les entités Select :
     - Articles : choisissez d'explorer les articles, leurs commentaires, leurs commentaires et leurs pièces jointes.
     - Pages de base : choisissez si vous souhaitez explorer les pages de base, leurs commentaires et leurs pièces jointes.
     - Blocs de base : choisissez d'explorer les blocs de base, leurs commentaires et leurs pièces jointes.
     - Vous pouvez également choisir d'ajouter des types de contenu personnalisés et des blocs personnalisés.

   b. Pour une configuration supplémentaire, optionnelle :

   - Pour le modèle Regex : ajoutez des modèles d'expressions régulières pour inclure ou exclure des titres d'entités et des noms de fichiers spécifiques. Vous pouvez ajouter jusqu'à 100 motifs.

   c. Mode de synchronisation : choisissez la manière dont vous souhaitez mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous

synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation.

- Synchronisation complète : Indexez tout le contenu fraîchement, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
- Synchronisation nouvelle, modifiée ou supprimée : Indexez uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

d. Dans Synchroniser le calendrier d'exécution, Fréquence : fréquence Amazon Kendra de synchronisation avec votre source de données.

e. Choisissez Suivant.

8. Sur la page Définir les mappages de champs, entrez les informations suivantes :

a. Pour le contenu, les commentaires et les pièces jointes : sélectionnez parmi les champs de source de données par défaut Amazon Kendra générés que vous souhaitez mapper à votre index.

b. Ajouter un champ —Pour ajouter des champs de source de données personnalisés afin de créer un nom de champ d'index à mapper et le type de données du champ.

c. Choisissez Suivant.

9. Sur la page Réviser et créer, vérifiez que les informations saisies sont correctes, puis sélectionnez Ajouter une source de données. Vous pouvez également choisir de modifier vos informations depuis cette page. Votre source de données apparaîtra sur la page Sources de données une fois que la source de données aura été ajoutée avec succès.

API

Pour vous connecter Amazon Kendra à Drupal

Vous devez spécifier un JSON du schéma de source de données à l'aide de l'TemplateConfigurationAPI. Vous devez fournir les informations suivantes :

- Source de données —Spécifiez le type de source de données tel que `DRUPAL` lorsque vous utilisez le schéma TemplateConfigurationJSON. Spécifiez également la source de données `TEMPLATE` lorsque vous appelez l'CreateDataSourceAPI.
- Mode de synchronisation : spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation. Vous pouvez choisir entre :
  - `FORCED_FULL_CRAWL`pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
  - `FULL_CRAWL`pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
  - `CHANGE_LOG`pour indexer uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
- Nom de ressource Amazon secret (ARN) : fournissez le nom de ressource Amazon (ARN) d'un Secrets Manager secret contenant les informations d'authentification que vous avez créées dans votre compte Drupal.

  Si vous utilisez l'authentification de base, le secret est stocké dans une structure JSON avec les clés suivantes :

  ```
  {
      "username": "user name",
      "password": "password"
  }
  ```

  Si vous utilisez l'authentification OAuth 2.0, le secret est stocké dans une structure JSON avec les clés suivantes :

  ```
  {
      "username": "user name",
      "password": "password",
  ```

```
    "clientId": "client id",
    "clientSecret": "client secret"
}
```

> **Note**
>
> > **Note**
> >
> > Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

- IAM role —Spécifiez `RoleArn` quand vous appelez `CreateDataSource` pour fournir à un IAM rôle les autorisations d'accéder à votre Secrets Manager secret et d'appeler le public requis APIs pour le connecteur Drupal et. Amazon Kendra Pour plus d'informations, consultez la section IAM Rôles des sources de données Drupal.

Vous pouvez également ajouter les fonctionnalités optionnelles suivantes :

- Virtual Private Cloud (VPC) **`VpcConfiguration`** —Spécifiez le moment de votre appel. `CreateDataSource` Pour de plus amples informations, veuillez consulter Configuration Amazon Kendra pour utiliser un Amazon VPC.
- Filtres d'inclusion et d'exclusion : vous pouvez indiquer si vous souhaitez inclure le contenu, les commentaires et les pièces jointes. Vous pouvez également spécifier des modèles d'expressions régulières pour inclure ou exclure du contenu, des commentaires et des pièces jointes.

> **Note**
>
> La plupart des sources de données utilisent des modèles d'expressions régulières, qui sont des modèles d'inclusion ou d'exclusion appelés filtres. Si vous spécifiez un filtre d'inclusion, seul le contenu correspondant au filtre d'inclusion est indexé. Tout document qui ne correspond pas au filtre d'inclusion n'est pas indexé. Si vous spécifiez

un filtre d'inclusion et d'exclusion, les documents correspondant au filtre d'exclusion ne sont pas indexés, même s'ils correspondent au filtre d'inclusion.

- Identity Crawler : spécifiez s'il faut activer l'explorateur Amazon Kendra d'identité. Le robot d'exploration d'identité utilise les informations de la liste de contrôle d'accès (ACL) de vos documents pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Si vous disposez d'une ACL pour vos documents et que vous choisissez de l'utiliser, vous pouvez également choisir d'activer le robot d'exploration Amazon Kendra d'identité pour configurer le filtrage contextuel utilisateur des résultats de recherche. Sinon, si le robot d'identification est désactivé, tous les documents peuvent être consultés publiquement. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'exploration d'identité est désactivé, vous pouvez également utiliser l'PutPrincipalMappingAPI pour télécharger les informations d'accès des utilisateurs et des groupes afin de filtrer le contexte utilisateur.
- Mappages de champs : choisissez de mapper les champs de votre source de données Drupal à vos champs d'index. Amazon Kendra Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.

> **Note**
>
> Le champ du corps du document ou l'équivalent du corps du document pour vos documents est requis pour Amazon Kendra effectuer une recherche dans vos documents. Vous devez associer le nom du champ du corps du document dans votre source de données au nom du champ d'index`_document_body`. Tous les autres champs sont facultatifs.

Pour obtenir la liste des autres clés JSON importantes à configurer, consultezSchéma du modèle Drupal.

Remarques

- Drupal n'a APIs pas de limite officielle de régulation.
- Java SDKs n'est pas disponible pour Drupal.
- Les données Drupal ne peuvent être récupérées qu'à l'aide des API JSON natives.
- Les types de contenu qui ne sont associés à aucune vue Drupal ne peuvent pas être analysés.

- Vous devez disposer d'un accès administrateur pour analyser les données de Drupal Blocks.
- Aucune API JSON n'est disponible pour créer le type de contenu défini par l'utilisateur à l'aide de verbes HTTP.
- Le corps du document et les commentaires relatifs aux articles, aux pages de base, aux blocs de base, au type de contenu défini par l'utilisateur et au type de bloc défini par l'utilisateur sont affichés au format HTML. Si le contenu HTML n'est pas bien formé, les balises associées au HTML apparaîtront dans le corps du document et dans les commentaires et seront visibles dans les résultats Amazon Kendra de recherche.
- Les types de contenu et les types de blocs sans description ni corps ne seront pas ingérés. Amazon Kendra Seuls les commentaires et les pièces jointes de ce type de contenu ou de bloc seront ingérés dans votre Amazon Kendra index.

GitHub

GitHub est un service d'hébergement Web pour le développement de logiciels fournissant des services de stockage et de gestion de code avec contrôle de version. Vous pouvez les utiliser Amazon Kendra pour indexer les fichiers de vos référentiels GitHub GitHub Enterprise Cloud (SaaS) et Enterprise Server (On Prem), émettre et extraire des demandes, émettre et extraire des commentaires, et émettre et extraire des pièces jointes aux commentaires. Vous pouvez également choisir d'inclure ou d'exclure certains fichiers.

> **Note**
>
> Amazon Kendra prend désormais en charge un GitHub connecteur amélioré.
> La console a été automatiquement mise à niveau pour vous. Tous les nouveaux connecteurs que vous créez dans la console utiliseront l'architecture mise à niveau. Si vous utilisez l'API, vous devez désormais utiliser TemplateConfigurationobjet au lieu de l'`GitHubConfiguration`objet pour configurer votre connecteur.
> Les connecteurs configurés à l'aide de l'ancienne console et de l'ancienne architecture d'API continueront de fonctionner tels qu'ils ont été configurés. Toutefois, vous ne pourrez ni les modifier ni les mettre à jour. Si vous souhaitez modifier ou mettre à jour la configuration de votre connecteur, vous devez créer un nouveau connecteur.
> Nous vous recommandons de migrer le flux de travail de votre connecteur vers la version mise à niveau. Support pour les connecteurs configurés à l'aide de l'ancienne architecture devrait prendre fin en juin 2024.

Vous pouvez vous connecter Amazon Kendra à votre source de GitHub données à l'aide de la Amazon Kendra console et de l'TemplateConfigurationAPI.

Pour résoudre les problèmes liés à votre connecteur de source de Amazon Kendra GitHub données, consultezDépannage des sources de données.

Rubriques

- Fonctionnalités prises en charge
- Prérequis
- Instructions de connexion
- En savoir plus

Fonctionnalités prises en charge

Amazon Kendra GitHub le connecteur de source de données prend en charge les fonctionnalités suivantes :

- Mappages de champs
- Contrôle d'accès des utilisateurs
- Filtres d'inclusion/exclusion
- Synchronisation complète et incrémentielle du contenu
- Cloud privé virtuel (VPC)

Prérequis

Avant de pouvoir Amazon Kendra indexer votre source de GitHub données, apportez ces modifications à vos AWS comptes GitHub and.

Dans GitHub, assurez-vous d'avoir :

- A créé un GitHub utilisateur doté d'autorisations administratives pour l' GitHub organisation.
- Vous avez configuré un jeton d'accès personnel dans Git Hub à utiliser comme identifiant d'authentification. Consultez GitHub la documentation sur la création d'un jeton d'accès personnel.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

- Recommandé : j'ai configuré un OAuth jeton pour les informations d'authentification. Utilisez un OAuth jeton pour améliorer les limites d'accélération des API et les performances du connecteur. Voir GitHub la documentation sur OAuth l'autorisation.
- Notez l'URL de l' GitHub hôte correspondant au type de GitHub service que vous utilisez. Par exemple, l'URL de l'hôte pour le GitHub cloud peut être *`https://api.github.com`* et l'URL de l'hôte pour le GitHub serveur peut être*`https://on-prem-host-url/api/v3/`*.
- Notez le nom de votre organisation pour GitHub le compte GitHub Enterprise Cloud (SaaS) ou le compte GitHub Enterprise Server (sur site) auquel vous souhaitez vous connecter. Vous pouvez trouver le nom de votre organisation en vous connectant à votre GitHub ordinateur de bureau et en sélectionnant Vos organisations dans le menu déroulant de votre photo de profil.
- Facultatif (serveur uniquement) : création d'un certificat SSL et copie du chemin d'accès au certificat stocké dans un Amazon S3 compartiment. Vous pouvez l'utiliser pour vous connecter GitHub si vous avez besoin d'une connexion SSL sécurisée. Vous pouvez simplement générer un certificat X509 auto-signé sur n'importe quel ordinateur utilisant OpenSSL. Pour un exemple d'utilisation d'OpenSSL pour créer un certificat X509, voir Créer et signer un certificat X509.
- Les autorisations suivantes ont été ajoutées :

  Pour le cloud GitHub d'entreprise (SaaS)

  - `repo:status`— Accorde un accès en lecture/écriture aux statuts de validation dans les référentiels publics et privés. Cette étendue est uniquement nécessaire pour accorder à d'autres utilisateurs ou services l'accès aux statuts de validation du référentiel privé sans accorder l'accès au code.
  - `repo_deployment`— Accorde l'accès aux statuts de déploiement pour les référentiels publics et privés. Cette étendue est uniquement nécessaire pour accorder à d'autres utilisateurs ou services l'accès aux statuts de déploiement, sans accorder l'accès au code.
  - `public_repo`— Limite l'accès aux référentiels publics. Cela inclut l'accès en lecture/écriture au code, les statuts de validation, les projets de référentiel, les collaborateurs et les statuts de

déploiement pour les référentiels publics et les organisations. Également nécessaire pour mettre en vedette les référentiels publics.

- `repo:invite`— Accorde des capacités d'acceptation/de refus pour les invitations à collaborer sur un référentiel. Cette étendue est uniquement nécessaire pour accorder à d'autres utilisateurs ou services l'accès aux invitations sans accorder l'accès au code.
- `security_events`— Accords : accès en lecture et en écriture aux événements de sécurité dans l'API de numérisation de code. Cette étendue est uniquement nécessaire pour autoriser d'autres utilisateurs ou services à accéder aux événements de sécurité sans accorder l'accès au code.
- `read:org`— Accès en lecture seule aux membres de l'organisation, aux projets de l'organisation et aux membres de l'équipe.
- `user:email`— Accorde un accès en lecture aux adresses e-mail des utilisateurs. Requis par Amazon Kendra pour explorer. ACLs
- `user:follow`— Permet de suivre ou de ne plus suivre d'autres utilisateurs. Requis par Amazon Kendra pour explorer. ACLs
- `read:user`— Autorise l'accès pour lire les données du profil d'un utilisateur. Requis par Amazon Kendra pour explorer. ACLs
- `workflow`— Permet d'ajouter et de mettre à jour des fichiers de flux de travail GitHub Actions. Les fichiers de flux de travail peuvent être validés sans cette étendue si le même fichier (avec le même chemin et le même contenu) existe sur une autre branche du même référentiel.

Pour plus d'informations, voir Étendue des OAuth applications dans GitHub Documentations.

Pour GitHub Enterprise Server (sur site)

- `repo:status`— Accorde un accès en lecture/écriture aux statuts de validation dans les référentiels publics et privés. Cette étendue est uniquement nécessaire pour accorder à d'autres utilisateurs ou services l'accès aux statuts de validation du référentiel privé sans accorder l'accès au code.
- `repo_deployment`— Accorde l'accès aux statuts de déploiement pour les référentiels publics et privés. Cette étendue est uniquement nécessaire pour accorder à d'autres utilisateurs ou services l'accès aux statuts de déploiement, sans accorder l'accès au code.
- `public_repo`— Limite l'accès aux référentiels publics. Cela inclut l'accès en lecture/écriture au code, les statuts de validation, les projets de référentiel, les collaborateurs et les statuts de déploiement pour les référentiels publics et les organisations. Également nécessaire pour mettre en vedette les référentiels publics.

- `repo:invite`— Accorde des capacités d'acceptation/de refus pour les invitations à collaborer sur un référentiel. Cette étendue est uniquement nécessaire pour accorder à d'autres utilisateurs ou services l'accès aux invitations sans accorder l'accès au code.
- `security_events`— Accords : accès en lecture et en écriture aux événements de sécurité dans l'API de numérisation de code. Cette étendue est uniquement nécessaire pour autoriser d'autres utilisateurs ou services à accéder aux événements de sécurité sans accorder l'accès au code.
- `read:user`— Autorise l'accès pour lire les données du profil d'un utilisateur. Requis par Amazon Q Business pour l'exploration ACLs.
- `user:email`— Accorde un accès en lecture aux adresses e-mail des utilisateurs. Requis par Amazon Q Business pour l'exploration ACLs.
- `user:follow`— Permet de suivre ou de ne plus suivre d'autres utilisateurs. Requis par Amazon Q Business pour l'exploration ACLs.
- `site_admin`— Accorde aux administrateurs du site l'accès aux points de terminaison de l'API d'administration des serveurs d' GitHub entreprise.
- `workflow`— Permet d'ajouter et de mettre à jour des fichiers de flux de travail GitHub Actions. Les fichiers de flux de travail peuvent être validés sans cette étendue si le même fichier (avec le même chemin et le même contenu) existe sur une autre branche du même référentiel.

  Pour plus d'informations, voir Étendue des OAuth applications dans GitHub Champs de documentation et de compréhension pour les OAuth applications dans GitHub Développeur.

- Il est vérifié que chaque document est unique dans GitHub et entre les autres sources de données que vous prévoyez d'utiliser pour le même index. Chaque source de données que vous souhaitez utiliser pour un index ne doit pas contenir le même document dans toutes les sources de données. IDs Les documents sont globaux par rapport à un index et doivent être uniques par index.

Dans votre Compte AWS, assurez-vous d'avoir :

- Création d'un Amazon Kendra index et, si vous utilisez l'API, notez l'ID de l'index.
- Vous avez créé un IAM rôle pour votre source de données et, si vous utilisez l'API, notez l'ARN du IAM rôle.

> **Note**
>
> Si vous modifiez votre type d'authentification et vos informations d'identification, vous devez mettre à jour votre IAM rôle pour accéder au bon identifiant AWS Secrets Manager secret.

- Stockez vos informations d' GitHub authentification dans un AWS Secrets Manager secret et, si vous utilisez l'API, notez l'ARN du secret.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

Si vous n'avez pas de IAM rôle ou de secret existant, vous pouvez utiliser la console pour créer un nouveau IAM rôle et un nouveau Secrets Manager secret lorsque vous connectez votre source de GitHub données à Amazon Kendra. Si vous utilisez l'API, vous devez fournir l'ARN d'un IAM rôle et d'un Secrets Manager secret existants, ainsi qu'un identifiant d'index.

Instructions de connexion

Pour vous connecter Amazon Kendra à votre source de GitHub données, vous devez fournir les informations nécessaires sur votre source de GitHub données afin de Amazon Kendra pouvoir accéder à vos données. Si vous n'avez pas encore configuré GitHub pour Amazon Kendra, consultezPrérequis.

Console

Pour vous connecter Amazon Kendra à GitHub

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la.
2. Dans le volet de navigation de gauche, choisissez Index, puis choisissez l'index que vous souhaitez utiliser dans la liste des index.

> **Note**
>
> Vous pouvez choisir de configurer ou de modifier vos paramètres de contrôle d'accès utilisateur dans les paramètres de l'index.

3. Sur la page de démarrage, choisissez Ajouter une source de données.
4. Sur la page Ajouter une source de données, choisissez GitHub connecteur, puis sélectionnez Ajouter un connecteur. Si vous utilisez la version 2 (le cas échéant), choisissez le GitHub connecteur avec la balise « V2.0 ».
5. Sur la page Spécifier les détails de la source de données, entrez les informations suivantes :
   a. Dans Nom et description, pour Nom de la source de données : entrez le nom de votre source de données. Vous pouvez inclure des traits d'union, mais pas des espaces.
   b. (Facultatif) Description : entrez une description facultative pour votre source de données.
   c. Dans la langue par défaut : choisissez une langue pour filtrer vos documents pour l'index. Sauf indication contraire, la langue par défaut est l'anglais. La langue spécifiée dans les métadonnées du document remplace la langue sélectionnée.
   d. Dans Balises, pour Ajouter une nouvelle balise : incluez des balises facultatives pour rechercher et filtrer vos ressources ou suivre vos AWS coûts.
   e. Choisissez Suivant.
6. Sur la page Définir l'accès et la sécurité, entrez les informations suivantes :
   a. GitHubsource —Choisissez entre GitHub Enterprise Cloud et GitHubEnterprise Server.
   b. GitHub URL de l'hôte —Par exemple, l'URL de l'hôte pour le GitHub cloud peut être *`https://api.github.com`* et l'URL de l'hôte pour le GitHub serveur peut être*`https://on-prem-host-url/api/v3/`*.
   c. GitHub nom de l'organisation : entrez le nom de votre GitHub organisation. Vous pouvez trouver les informations relatives à votre organisation dans votre GitHub compte.

   > **Note**
   >
   > GitHub connector prend en charge l'analyse d'une seule organisation par instance de connecteur de source de données.

d. Autorisation : activez ou désactivez les informations de la liste de contrôle d'accès (ACL) pour vos documents, si vous disposez d'une ACL et que vous souhaitez l'utiliser pour le contrôle d'accès. L'ACL indique les documents auxquels les utilisateurs et les groupes peuvent accéder. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.

e. AWS Secrets Manager secret —Choisissez un secret existant ou créez-en un nouveau Secrets Manager pour stocker vos informations GitHub d'authentification. Si vous choisissez de créer un nouveau secret, une fenêtre AWS Secrets Manager secrète s'ouvre.

   i. Entrez les informations suivantes dans la fenêtre Créer un AWS Secrets Manager secret :

      A. Nom secret : le nom de votre secret. Le préfixe « AmazonKendra - GitHub - » est automatiquement ajouté à votre nom secret.

      B. Pour le GitHubjeton : entrez la valeur d'identification d'authentification configurée dans. GitHub

   ii. Enregistrez et ajoutez votre secret.

f. Virtual Private Cloud (VPC) —Vous pouvez choisir d'utiliser un VPC. Dans ce cas, vous devez ajouter des sous-réseaux et des groupes de sécurité VPC.

g. Identity Crawler : spécifiez s'il faut activer l'explorateur Amazon Kendra d'identité. Le robot d'exploration d'identité utilise les informations de la liste de contrôle d'accès (ACL) de vos documents pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Si vous disposez d'une ACL pour vos documents et que vous choisissez de l'utiliser, vous pouvez également choisir d'activer le robot d'exploration Amazon Kendra d'identité pour configurer le filtrage contextuel utilisateur des résultats de recherche. Sinon, si le robot d'identification est désactivé, tous les documents peuvent être consultés publiquement. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'exploration d'identité est désactivé, vous pouvez également utiliser l'PutPrincipalMappingAPI pour télécharger les informations d'accès des utilisateurs et des groupes afin de filtrer le contexte utilisateur.

h. IAM rôle —Choisissez un IAM rôle existant ou créez-en un nouveau IAM pour accéder aux informations d'identification de votre référentiel et indexer le contenu.

> **Note**
>
> IAM les rôles utilisés pour les index ne peuvent pas être utilisés pour les sources de données. Si vous ne savez pas si un rôle existant est utilisé pour un index ou une FAQ, choisissez Créer un nouveau rôle pour éviter les erreurs.

i. Choisissez Suivant.

7. Sur la page Configurer les paramètres de synchronisation, entrez les informations suivantes :

   a. Sélectionnez les référentiels : choisissez d'explorer tous les référentiels ou de sélectionner.

      Si vous choisissez d'explorer certains référentiels, ajoutez les noms des référentiels et, éventuellement, le nom de branches spécifiques.

   b. Types de contenu : choisissez les types de contenu que vous souhaitez explorer parmi les fichiers, les problèmes, les pull requests, etc.

   c. Modèles Regex : ajoutez des modèles d'expressions régulières pour inclure ou exclure certains fichiers.

   d. Mode de synchronisation : choisissez la manière dont vous souhaitez mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation.

      - Synchronisation complète : Indexez tout le contenu fraîchement, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
      - Nouvelle synchronisation modifiée : Indexez uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
      - Synchronisation nouvelle, modifiée ou supprimée : Indexez uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre

source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

e. Dans Synchroniser le calendrier d'exécution pour la fréquence : choisissez la fréquence à laquelle vous souhaitez synchroniser le contenu de votre source de données et mettre à jour votre index.

f. Choisissez Suivant.

8. Sur la page Définir les mappages de champs, entrez les informations suivantes :

   a. Champs de source de données par défaut : sélectionnez parmi les champs de source de données par défaut Amazon Kendra générés que vous souhaitez mapper à votre index.

   b. Ajouter un champ —Pour ajouter des champs de source de données personnalisés afin de créer un nom de champ d'index à mapper et le type de données du champ.

   c. Choisissez Suivant.

9. Sur la page Réviser et créer, vérifiez que les informations saisies sont correctes, puis sélectionnez Ajouter une source de données. Vous pouvez également choisir de modifier vos informations à partir de cette page. Votre source de données apparaîtra sur la page Sources de données une fois que la source de données aura été ajoutée avec succès.

API

Pour vous connecter Amazon Kendra à GitHub

Vous devez spécifier un code JSON du schéma de source de données à l'aide du TemplateConfigurationAPI. Vous devez fournir les informations suivantes :

- Source de données —Spécifiez le type de source de données comme `GITHUB` lorsque vous utilisez TemplateConfigurationSchéma JSON. Spécifiez également la source de données comme `TEMPLATE` lorsque vous appelez le CreateDataSourceAPI.
- GitHubtype —Spécifiez le type comme suit : `SAAS` ou`ON_PREMISE`.
- URL de l'hôte : spécifiez l'URL de l' GitHub hôte ou l'URL du point de terminaison de l'API. Par exemple, si vous utilisez le GitHub SaaS/Enterprise Cloud, l'URL de l'hôte peut être, et pour le serveur GitHub local ou d'entreprise`https://api.github.com`, l'URL de l'hôte peut être l'URL de l'hôte. `https://on-prem-host-url/api/v3/`
- Nom de l'organisation —Spécifiez le nom de l'organisation du GitHub compte. Vous pouvez trouver le nom de votre organisation en vous connectant à votre GitHub ordinateur de bureau et en sélectionnant Vos organisations dans le menu déroulant de votre photo de profil.

- Mode de synchronisation : spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation. Vous pouvez choisir entre :
  - `FORCED_FULL_CRAWL`pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
  - `FULL_CRAWL`pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
  - `CHANGE_LOG`pour indexer uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
- Identity Crawler : spécifiez s'il faut activer l'explorateur Amazon Kendra d'identité. Le robot d'exploration d'identité utilise les informations de la liste de contrôle d'accès (ACL) de vos documents pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Si vous disposez d'une ACL pour vos documents et que vous choisissez de l'utiliser, vous pouvez également choisir d'activer le robot d'exploration Amazon Kendra d'identité pour configurer le filtrage contextuel utilisateur des résultats de recherche. Sinon, si le robot d'identification est désactivé, tous les documents peuvent être consultés publiquement. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'exploration d'identité est désactivé, vous pouvez également utiliser l'PutPrincipalMappingAPI pour télécharger les informations d'accès des utilisateurs et des groupes afin de filtrer le contexte utilisateur.
- Nom de ressource Amazon (ARN) secret : indiquez le nom de ressource Amazon (ARN) d'un Secrets Manager secret contenant les informations d'authentification de votre GitHub compte. Le secret est stocké dans une structure JSON avec les clés suivantes :

```
{
    "personalToken": "token"
}
```

- IAM role —Spécifiez le `RoleArn` moment où vous appelez `CreateDataSource` pour fournir à un IAM rôle les autorisations d'accéder à votre Secrets Manager secret et d'appeler le public requis APIs pour le GitHub connecteur et Amazon Kendra. Pour plus d'informations, consultez la section IAM Rôles des sources de GitHub données.

Vous pouvez également ajouter les fonctionnalités optionnelles suivantes :

- Virtual Private Cloud (VPC) **`VpcConfiguration`** —Spécifiez le moment de votre appel. `CreateDataSource` Pour de plus amples informations, veuillez consulter Configuration Amazon Kendra pour utiliser un Amazon VPC.

> **Note**
>
> Si vous utilisez un GitHub serveur, vous devez utiliser un Amazon VPC pour vous connecter à votre GitHub serveur.

- Filtre de référentiel : filtrez les référentiels en fonction de leur nom et du nom de leurs branches.
- Types de documents/contenus : indiquez s'il faut analyser les documents du référentiel, les problèmes, les commentaires, les pièces jointes aux commentaires, les pull requests, les commentaires des demandes d'extraction, les pièces jointes aux commentaires des demandes d'extraction.
- Filtres d'inclusion et d'exclusion : indiquez si vous souhaitez inclure ou exclure certains fichiers et dossiers.

> **Note**
>
> La plupart des sources de données utilisent des modèles d'expressions régulières, qui sont des modèles d'inclusion ou d'exclusion appelés filtres. Si vous spécifiez un filtre d'inclusion, seul le contenu correspondant au filtre d'inclusion est indexé. Tout document qui ne correspond pas au filtre d'inclusion n'est pas indexé. Si vous spécifiez un filtre d'inclusion et d'exclusion, les documents correspondant au filtre d'exclusion ne sont pas indexés, même s'ils correspondent au filtre d'inclusion.

- Liste de contrôle d'accès (ACL) : indiquez si vous souhaitez analyser les informations ACL de vos documents, si vous disposez d'une ACL et souhaitez l'utiliser pour le contrôle d'accès. L'ACL indique les documents auxquels les utilisateurs et les groupes peuvent accéder. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de

l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.

- Mappages de champs : choisissez de mapper les champs de votre source de GitHub données à vos champs d' Amazon Kendra index. Vous pouvez inclure des champs de documents, des validations, des problèmes, des pièces jointes, des commentaires, des pull requests, des pièces jointes de pull request, des commentaires de pull request. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.

> **Note**
>
> Le champ du corps du document ou l'équivalent du corps du document pour vos documents est requis pour qu'Amazon Kendra puisse effectuer des recherches dans vos documents. Vous devez associer le nom du champ du corps du document dans votre source de données au nom du champ d'index`_document_body`. Tous les autres champs sont facultatifs.

Pour obtenir la liste des autres clés JSON importantes à configurer, voir GitHub schéma de modèle.

En savoir plus

Pour en savoir plus sur l'intégration Amazon Kendra à votre source de GitHub données, consultez :

- Réimaginez la recherche sur GitHub les référentiels grâce à la puissance du connecteur Amazon Kendra GitHub

Gmail

Gmail est un client de messagerie développé par Google grâce auquel vous pouvez envoyer des e-mails avec des pièces jointes. Les messages Gmail peuvent être triés et stockés dans votre boîte de réception à l'aide de dossiers et d'étiquettes. Vous pouvez l'utiliser Amazon Kendra pour indexer vos e-mails et vos pièces jointes. Vous pouvez également configurer Amazon Kendra pour inclure ou exclure des e-mails, des pièces jointes et des étiquettes spécifiques à des fins d'indexation.

Vous pouvez vous connecter Amazon Kendra à votre source de données Gmail à l'aide de la Amazon Kendra console et de l'TemplateConfigurationAPI.

Pour résoudre les problèmes liés à votre connecteur de source de données Amazon Kendra Gmail, consultezDépannage des sources de données.

Rubriques

- Fonctionnalités prises en charge
- Prérequis
- Instructions de connexion
- En savoir plus
- Remarques

Fonctionnalités prises en charge

- Mappages de champs
- Contrôle d'accès des utilisateurs
- Filtres d'inclusion/exclusion
- Synchronisation complète et incrémentielle du contenu
- Cloud privé virtuel (VPC)

Prérequis

Avant de pouvoir Amazon Kendra indexer votre source de données Gmail, apportez ces modifications à votre compte Gmail et à vos AWS comptes.

Dans Gmail, assurez-vous que vous disposez des éléments suivants :

- J'ai créé un compte administrateur Google Cloud Platform et j'ai créé un projet Google Cloud.
- Vous avez activé l'API Gmail et l'API Admin SDK dans votre compte administrateur.
- Vous avez créé un compte de service et téléchargé une clé privée JSON pour votre compte Gmail. Pour plus d'informations sur la création et l'accès à votre clé privée, consultez la documentation de Google Cloud sur la création d'une clé de compte de service et les informations d'identification du compte de service.
- Vous avez copié l'adresse e-mail de votre compte administrateur, celle de votre compte de service et votre clé privée pour les utiliser comme informations d'authentification.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

- Les étendues Oauth suivantes (en utilisant un rôle d'administrateur) ont été ajoutées pour votre utilisateur et les répertoires partagés que vous souhaitez indexer :
  - https://www.googleapis.com/auth/admin.directory.user.readonly
  - https://www.googleapis.com/auth/gmail.readonly
- Il est vérifié que chaque document est unique dans Gmail et dans les autres sources de données que vous prévoyez d'utiliser pour le même index. Chaque source de données que vous souhaitez utiliser pour un index ne doit pas contenir le même document dans toutes les sources de données. IDs Les documents sont globaux par rapport à un index et doivent être uniques par index.

Dans votre Compte AWS, assurez-vous d'avoir :

- Création d'un Amazon Kendra index et, si vous utilisez l'API, notez l'ID de l'index.
- Vous avez créé un IAM rôle pour votre source de données et, si vous utilisez l'API, notez l'ARN du IAM rôle.

> **Note**
>
> Si vous modifiez votre type d'authentification et vos informations d'identification, vous devez mettre à jour votre IAM rôle pour accéder au bon identifiant AWS Secrets Manager secret.

- Stockez vos informations d'authentification Gmail dans un AWS Secrets Manager secret et, si vous utilisez l'API, notez l'ARN du secret.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire

pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

Si vous n'avez pas de IAM rôle ou de secret existant, vous pouvez utiliser la console pour créer un nouveau IAM rôle et un nouveau Secrets Manager secret lorsque vous connectez votre source de données Gmail à Amazon Kendra. Si vous utilisez l'API, vous devez fournir l'ARN d'un IAM rôle et d'un Secrets Manager secret existants, ainsi qu'un identifiant d'index.

Instructions de connexion

Pour vous connecter Amazon Kendra à votre source de données Gmail, vous devez fournir les détails de vos informations d'identification Gmail afin de Amazon Kendra pouvoir accéder à vos données. Si vous n'avez pas encore configuré Gmail pour Amazon Kendra, consultezPrérequis.

Console

Pour vous connecter Amazon Kendra à Gmail

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la.
2. Dans le volet de navigation de gauche, choisissez Index, puis choisissez l'index que vous souhaitez utiliser dans la liste des index.

   > **Note**
   >
   > Vous pouvez choisir de configurer ou de modifier vos paramètres de contrôle d'accès utilisateur dans les paramètres de l'index.

3. Sur la page de démarrage, choisissez Ajouter une source de données.
4. Sur la page Ajouter une source de données, choisissez le connecteur Gmail, puis sélectionnez Ajouter un connecteur. Si vous utilisez la version 2 (le cas échéant), choisissez le connecteur Gmail avec la balise « V2.0 ».
5. Sur la page Spécifier les détails de la source de données, entrez les informations suivantes :

   a. Dans Nom et description, pour Nom de la source de données : entrez le nom de votre source de données. Vous pouvez inclure des traits d'union, mais pas des espaces.

   b. (Facultatif) Description : entrez une description facultative pour votre source de données.

c. Dans la langue par défaut : choisissez une langue pour filtrer vos documents pour l'index. Sauf indication contraire, la langue par défaut est l'anglais. La langue spécifiée dans les métadonnées du document remplace la langue sélectionnée.

d. Dans Balises, pour Ajouter une nouvelle balise : incluez des balises facultatives pour rechercher et filtrer vos ressources ou suivre vos AWS coûts.

e. Choisissez Suivant.

6. Sur la page Définir l'accès et la sécurité, entrez les informations suivantes :

a. Autorisation : activez ou désactivez les informations de la liste de contrôle d'accès (ACL) pour vos documents, si vous disposez d'une ACL et que vous souhaitez l'utiliser pour le contrôle d'accès. L'ACL indique les documents auxquels les utilisateurs et les groupes peuvent accéder. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.

b. Dans Authentification pour le AWS Secrets Manager secret : choisissez un secret existant ou créez-en un nouveau Secrets Manager pour stocker vos informations d'authentification Gmail. Si vous choisissez de créer un nouveau secret, une fenêtre AWS Secrets Manager secrète s'ouvre.

- Entrez les informations suivantes dans la fenêtre Créer un AWS Secrets Manager secret :

A. Nom secret : le nom de votre secret.

B. Adresse e-mail du client : adresse e-mail du client que vous avez copiée depuis votre compte de service Google.

C. Adresse e-mail du compte administrateur : adresse e-mail du compte administrateur que vous souhaitez utiliser.

D. Clé privée : clé privée que vous avez copiée depuis votre compte de service Google.

E. Enregistrez et ajoutez votre secret.

c. Virtual Private Cloud (VPC) —Vous pouvez choisir d'utiliser un VPC. Dans ce cas, vous devez ajouter des sous-réseaux et des groupes de sécurité VPC.

d. IAM rôle —Choisissez un IAM rôle existant ou créez-en un nouveau IAM pour accéder aux informations d'identification de votre référentiel et indexer le contenu.

> **Note**
>
> IAM les rôles utilisés pour les index ne peuvent pas être utilisés pour les sources de données. Si vous ne savez pas si un rôle existant est utilisé pour un index ou une FAQ, choisissez Créer un nouveau rôle pour éviter les erreurs.

e. Choisissez Suivant.

7. Sur la page Configurer les paramètres de synchronisation, entrez les informations suivantes :

   a. Pour les types d'entités : choisissez de synchroniser les pièces jointes des messages.

   b. (Facultatif) Pour Configuration supplémentaire, entrez les informations suivantes :

      i. Plage de dates : entrez une plage de dates pour spécifier les dates de début et de fin des e-mails que vous souhaitez analyser.

      ii. Domaines de messagerie : incluez ou excluez certains e-mails en fonction des domaines de messagerie « à », « de », « cc » et « bcc ».

      iii. Mots clés dans les sujets : incluez ou excluez les e-mails en fonction des mots clés figurant dans leur objet.

      > **Note**
      >
      > Vous pouvez également choisir d'inclure tous les documents correspondant à tous les mots clés du sujet que vous avez saisis.

      iv. Étiquettes : ajoutez des modèles d'expressions régulières pour inclure ou exclure certaines étiquettes d'e-mail.

      v. Pièces jointes : ajoutez des modèles d'expressions régulières pour inclure ou exclure certaines pièces jointes à des e-mails.

   c. Mode de synchronisation : choisissez la manière dont vous souhaitez mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation.

- Synchronisation complète : Indexez tout le contenu fraîchement, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
- Synchronisation nouvelle, modifiée ou supprimée : Indexez uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

> **Important**
>
> Comme il n'existe aucune API permettant de mettre à jour les messages Gmail définitivement supprimés, les contenus nouveaux, modifiés ou supprimés sont synchronisés :
>
> - Ne supprimera pas de votre Amazon Kendra index les messages définitivement supprimés de Gmail
> - Ne synchronisera pas les modifications dans les libellés des e-mails Gmail
>
> Pour synchroniser les modifications apportées à l'étiquette de votre source de données Gmail et les e-mails définitivement supprimés avec votre Amazon Kendra index, vous devez effectuer régulièrement des analyses complètes.

d. Dans Calendrier d'exécution de la synchronisation, pour Fréquence : choisissez la fréquence à laquelle vous souhaitez synchroniser le contenu de votre source de données et mettre à jour votre index.

e. Choisissez Suivant.

8. Sur la page Définir les mappages de champs, entrez les informations suivantes :

a. Champs de source de données par défaut : sélectionnez parmi les champs de source de données par défaut Amazon Kendra générés que vous souhaitez mapper à votre index.

> **Note**
>
> Amazon Kendra Le connecteur de source de données Gmail ne prend pas en charge la création de champs d'index personnalisés en raison des limites de l'API.

b. Choisissez Suivant.

9. Sur la page Réviser et créer, vérifiez que les informations saisies sont correctes, puis sélectionnez Ajouter une source de données. Vous pouvez également choisir de modifier vos informations à partir de cette page. Votre source de données apparaîtra sur la page Sources de données une fois qu'elle aura été ajoutée avec succès.

API

Pour vous connecter Amazon Kendra à Gmail

Vous devez spécifier un code JSON du schéma de source de données à l'aide du TemplateConfigurationAPI. Vous devez fournir les informations suivantes :

- Source de données —Spécifiez le type de source de données `GMAIL` lorsque vous utilisez le TemplateConfigurationSchéma JSON. Spécifiez également la source de données comme `TEMPLATE` lorsque vous appelez le CreateDataSourceAPI.
- Mode de synchronisation : spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation. Vous pouvez choisir entre :
  - `FORCED_FULL_CRAWL`pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
  - `FULL_CRAWL`pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

> **Important**
>
> Comme il n'existe aucune API permettant de mettre à jour les messages Gmail définitivement supprimés, les contenus nouveaux, modifiés ou supprimés sont synchronisés :
>
> - Ne supprimera pas de votre Amazon Kendra index les messages définitivement supprimés de Gmail
> - Ne synchronisera pas les modifications dans les libellés des e-mails Gmail
>
> Pour synchroniser les modifications apportées à l'étiquette de votre source de données Gmail et les e-mails définitivement supprimés avec votre Amazon Kendra index, vous devez effectuer régulièrement des analyses complètes.

- Nom de ressource Amazon (ARN) secret : indiquez le nom de ressource Amazon (ARN) d'un Secrets Manager secret contenant les informations d'authentification de votre compte Gmail. Le secret est stocké dans une structure JSON avec les clés suivantes :

```
{
    "adminAccountEmailId": "service account email",
    "clientEmailId": "user account email",
    "privateKey": "private key"
}
```

- IAM role —Spécifiez à quel `RoleArn` moment vous appelez `CreateDataSource` pour fournir à un IAM rôle les autorisations d'accéder à votre Secrets Manager secret et d'appeler le public requis APIs pour le connecteur Gmail et Amazon Kendra. Pour plus d'informations, consultez la section IAM Rôles des sources de données Gmail.

Vous pouvez également ajouter les fonctionnalités optionnelles suivantes :

- Virtual Private Cloud (VPC) **`VpcConfiguration`** —Spécifiez le moment de votre appel. `CreateDataSource` Pour de plus amples informations, veuillez consulter Configuration Amazon Kendra pour utiliser un Amazon VPC.
- Filtres d'inclusion et d'exclusion : indiquez s'il faut inclure ou exclure certains e-mails « à », « de », « cc » ou « bcc ».

> **Note**
>
> La plupart des sources de données utilisent des modèles d'expressions régulières, qui sont des modèles d'inclusion ou d'exclusion appelés filtres. Si vous spécifiez un filtre d'inclusion, seul le contenu correspondant au filtre d'inclusion est indexé. Tout document qui ne correspond pas au filtre d'inclusion n'est pas indexé. Si vous spécifiez un filtre d'inclusion et d'exclusion, les documents correspondant au filtre d'exclusion ne sont pas indexés, même s'ils correspondent au filtre d'inclusion.

- Filtrage du contexte utilisateur et contrôle d'accèsAmazon Kendra : analyse la liste de contrôle d'accès (ACL) de vos documents, si vous disposez d'une ACL pour vos documents. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.
- Mappages de champs : choisissez de mapper les champs de votre source de données Gmail à vos champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.

> **Note**
>
> Le champ du corps du document ou l'équivalent du corps du document pour vos documents est requis pour Amazon Kendra effectuer une recherche dans vos documents. Vous devez associer le nom du champ du corps du document dans votre source de données au nom du champ d'index`_document_body`. Tous les autres champs sont facultatifs.

> **Note**
>
> Amazon Kendra Le connecteur de source de données Gmail ne prend pas en charge la création de champs d'index personnalisés en raison des limites de l'API.

Pour obtenir la liste des autres clés JSON importantes à configurer, voir Gmail schéma de modèle.

En savoir plus

Pour en savoir plus sur l'intégration Amazon Kendra à votre source de données Gmail, consultez :

- Effectuez une recherche intelligente dans les e-mails de votre espace de travail Google à l'aide du connecteur Gmail pour Amazon Kendra.

Remarques

- Comme il n'existe aucune API permettant de mettre à jour les messages Gmail définitivement supprimés, une synchronisation de contenu**`FULL_CRAWL`**/Nouveau, modifié ou supprimé :
  - Ne supprimera pas de votre Amazon Kendra index les messages définitivement supprimés de Gmail
  - Ne synchronisera pas les modifications dans les libellés des e-mails Gmail

  Pour synchroniser les modifications apportées à l'étiquette de votre source de données Gmail et les e-mails définitivement supprimés avec votre Amazon Kendra index, vous devez effectuer régulièrement des analyses complètes.
- Amazon Kendra Le connecteur de source de données Gmail ne prend pas en charge la création de champs d'index personnalisés en raison des limites de l'API.

Google Drive

Google Drive est un service de stockage de fichiers basé sur le cloud. Vous pouvez l'utiliser Amazon Kendra pour indexer les documents stockés dans des disques partagés, dans les dossiers Mes disques et dans les dossiers Shared with me de votre source de données Google Drive. Vous pouvez indexer à la fois les documents Google Workspace et les documents répertoriés dans Types de documentation. Vous pouvez également utiliser des filtres d'inclusion et d'exclusion pour indexer le contenu par nom de fichier, type de fichier et chemin de fichier.

Vous pouvez vous connecter Amazon Kendra à votre source de données Google Drive à l'aide de la Amazon Kendra console, de l'TemplateConfigurationAPI ou de l'GoogleDriveConfigurationAPI.

Amazon Kendra possède deux versions du connecteur Google Drive. Les fonctionnalités prises en charge par chaque version incluent :

Connecteur Google Drive V1.0/API GoogleDriveConfiguration

- Mappages de champs
- Contrôle d'accès utilisateur
- Filtres d'inclusion/exclusion

Connecteur Google Drive V2.0/API TemplateConfiguration

- Mappages de champs
- Contrôle d'accès utilisateur
- Filtres d'inclusion/exclusion
- Synchronisation complète et incrémentielle du contenu
- Cloud privé virtuel (VPC)

> **Note**
>
> Connecteur Google Drive V1.0/ L' DriveConfiguration API Google a pris fin en 2023. Nous vous recommandons de migrer vers ou d'utiliser le connecteur Google Drive TemplateConfiguration V2.0/API.

Pour résoudre les problèmes liés à votre connecteur de source de données Amazon Kendra Google Drive, consultezDépannage des sources de données.

Rubriques

- Connecteur Google Drive V1.0
- Connecteur Google Drive V2.0

Connecteur Google Drive V1.0

Google Drive est un service de stockage de fichiers basé sur le cloud. Vous pouvez l'utiliser Amazon Kendra pour indexer les documents et les commentaires stockés dans les dossiers Drive partagés, Mes Drives et Shared with me de votre source de données Google Drive. Vous pouvez indexer les documents Google Workspace, ainsi que les documents répertoriés dans la section Types de documentation. Vous pouvez également utiliser des filtres d'inclusion et d'exclusion pour indexer le contenu par nom de fichier, type de fichier et chemin de fichier.

> **Note**
>
> Connecteur Google Drive V1.0/ L' DriveConfiguration API Google a pris fin en 2023. Nous vous recommandons de migrer vers ou d'utiliser le connecteur Google Drive TemplateConfiguration V2.0/API.

Pour résoudre les problèmes liés à votre connecteur de source de données Amazon Kendra Google Drive, consultezDépannage des sources de données.

Rubriques

- Fonctionnalités prises en charge
- Prérequis
- Instructions de connexion
- En savoir plus

Fonctionnalités prises en charge

- Mappages de champs
- Contrôle d'accès utilisateur
- Filtres d'inclusion/exclusion

Prérequis

Avant de pouvoir Amazon Kendra indexer votre source de données Google Drive, apportez ces modifications à votre Google Drive et à vos AWS comptes.

Dans Google Drive, assurez-vous d'avoir :

- Vous avez obtenu l'accès par un rôle de super administrateur ou vous êtes un utilisateur doté de privilèges administratifs. Vous n'avez pas besoin d'un rôle de super administrateur si l'accès vous a été accordé par un rôle de super administrateur.
- Création d'un compte de service avec Activer la délégation à l'échelle du domaine G Suite activé et une clé JSON comme clé privée à l'aide du compte.
- Vous avez copié l'adresse e-mail de votre compte utilisateur et celle de votre compte de service. Lorsque vous vous connectez, entrez l'adresse e-mail de votre compte utilisateur en tant

qu'adresse e-mail du compte administrateur et l'adresse e-mail de votre compte de service en tant qu'e-mail client dans votre AWS Secrets Manager code secret. Amazon Kendra

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

- Ajout de l'API du SDK d'administration et de l'API Google Drive dans votre compte.
- Vous avez ajouté (ou demandé à un utilisateur doté d'un rôle de super administrateur d'ajouter) les autorisations suivantes à votre compte de service à l'aide d'un rôle de super administrateur :
  - https://www.googleapis.com/auth/lecteur.readonly
  - https://www.googleapis.com/auth/drive.metadata.readonly
  - https://www.googleapis.com/auth/admin.directory.user.readonly
  - https://www.googleapis.com/auth/admin.directory.group.readonly
- Il est vérifié que chaque document est unique dans Google Drive et dans les autres sources de données que vous prévoyez d'utiliser pour le même index. Chaque source de données que vous souhaitez utiliser pour un index ne doit pas contenir le même document dans toutes les sources de données. IDs Les documents sont globaux par rapport à un index et doivent être uniques par index.

Dans votre Compte AWS, assurez-vous d'avoir :

- Création d'un Amazon Kendra index et, si vous utilisez l'API, notez l'ID de l'index.
- Vous avez créé un IAM rôle pour votre source de données et, si vous utilisez l'API, notez l'ARN du IAM rôle.

> **Note**
>
> Si vous modifiez votre type d'authentification et vos informations d'identification, vous devez mettre à jour votre IAM rôle pour accéder au bon identifiant AWS Secrets Manager secret.

- Stockez vos informations d'authentification Google Drive dans un AWS Secrets Manager secret et, si vous utilisez l'API, notez l'ARN du secret.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

Si vous n'avez pas de IAM rôle ou de secret existant, vous pouvez utiliser la console pour créer un nouveau IAM rôle et un nouveau Secrets Manager secret lorsque vous connectez votre source de données Google Drive à Amazon Kendra. Si vous utilisez l'API, vous devez fournir l'ARN d'un IAM rôle et d'un Secrets Manager secret existants, ainsi qu'un identifiant d'index.

Instructions de connexion

Pour vous connecter Amazon Kendra à votre source de données Google Drive, vous devez fournir les informations nécessaires sur votre source de données Google Drive afin de Amazon Kendra pouvoir accéder à vos données. Si vous n'avez pas encore configuré Google Drive pour Amazon Kendra voirPrérequis.

Console

Pour vous connecter Amazon Kendra à Google Drive

1. Connectez-vous à la console AWS de gestion et Amazon Kendra ouvrez-la.
2. Dans le volet de navigation de gauche, choisissez Index, puis choisissez l'index que vous souhaitez utiliser dans la liste des index.

> **Note**
>
> Vous pouvez choisir de configurer ou de modifier vos paramètres de contrôle d'accès utilisateur dans les paramètres de l'index.

3. Sur la page de démarrage, choisissez Ajouter une source de données.

4. Sur la page Ajouter une source de données, choisissez le connecteur Google Drive V1.0, puis sélectionnez Ajouter un connecteur.
5. Sur la page Spécifier les détails de la source de données, entrez les informations suivantes :

   a. Dans Nom et description, pour Nom de la source de données : entrez le nom de votre source de données. Vous pouvez inclure des tirets, mais pas des espaces.

   b. (Facultatif) Description : entrez une description facultative pour votre source de données.

   c. Dans la langue par défaut : choisissez une langue pour filtrer vos documents pour l'index. Sauf indication contraire, la langue par défaut est l'anglais. La langue spécifiée dans les métadonnées du document remplace la langue sélectionnée.

   d. Dans Balises, pour Ajouter une nouvelle balise : incluez des balises facultatives pour rechercher et filtrer vos ressources ou suivre vos AWS coûts.

   e. Choisissez Suivant.
6. Sur la page Définir l'accès et la sécurité, entrez les informations suivantes :

   a. Pour le type d'authentification, choisissez entre existant et nouveau. Si vous choisissez d'utiliser un secret existant, utilisez Select secret pour choisir votre secret.

   b. Si vous choisissez de créer un nouveau secret, une option AWS Secrets Manager secrète s'ouvre.

      - Entrez les informations suivantes dans la fenêtre Créer un AWS Secrets Manager secret :

        A. Nom secret : nom de votre secret. Le préfixe « AmazonKendra -Google Drive » est automatiquement ajouté à votre nom secret.

        B. Pour l'adresse e-mail du compte administrateur, l'adresse e-mail du client et la clé privée, entrez les valeurs d'identification d'authentification que vous avez générées et téléchargées depuis votre compte Google Drive.

        C. Choisissez Enregistrer l'authentification.

   c. IAM rôle —Choisissez un IAM rôle existant ou créez-en un nouveau IAM pour accéder aux informations d'identification de votre référentiel et indexer le contenu.

> **Note**
>
> IAM les rôles utilisés pour les index ne peuvent pas être utilisés pour les sources de données. Si vous ne savez pas si un rôle existant est utilisé pour un index ou une FAQ, choisissez Créer un nouveau rôle pour éviter les erreurs.

d. Choisissez Suivant.

7. Sur la page Configurer les paramètres de synchronisation, entrez les informations suivantes :

   a. Exclure les comptes utilisateur : les utilisateurs de Google Drive que vous souhaitez exclure de l'index. Vous pouvez ajouter jusqu'à 100 comptes utilisateurs.

   b. Exclure les lecteurs partagés : les lecteurs partagés Google Drive que vous souhaitez exclure de votre index. Vous pouvez ajouter jusqu'à 100 disques partagés.

   c. Exclure les types de fichiers lecteurs : les types de fichiers Google Drive que vous souhaitez exclure de votre index. Vous pouvez également choisir de modifier les sélections de type MIME.

   d. Configurations supplémentaires : modèles d'expressions régulières pour inclure ou exclure certains contenus. Vous pouvez ajouter jusqu'à 100 motifs.

   e. Fréquence : fréquence Amazon Kendra de synchronisation avec votre source de données.

   f. Choisissez Suivant.

8. Sur la page Définir les mappages de champs, entrez les informations suivantes :

   a. Pour GoogleDrive le nom du champ et les mappages de champs supplémentaires suggérés, sélectionnez parmi les champs de source de données par défaut Amazon Kendra générés que vous souhaitez mapper à votre index.

   b. Ajouter un champ —Pour ajouter des champs de source de données personnalisés afin de créer un nom de champ d'index à mapper et le type de données du champ.

   c. Choisissez Suivant.

9. Sur la page Réviser et créer, vérifiez que les informations saisies sont correctes, puis sélectionnez Ajouter une source de données. Vous pouvez également choisir de modifier vos informations à partir de cette page. Votre source de données apparaîtra sur la page Sources de données une fois que la source de données aura été ajoutée avec succès.

API

Pour vous connecter Amazon Kendra à Google Drive

Vous devez spécifier les éléments suivants à l'aide de l'GoogleDriveConfigurationAPI :

- Nom de ressource Amazon secret (ARN) : indiquez le nom de ressource Amazon (ARN) d'un Secrets Manager secret contenant les informations d'authentification de votre compte Google Drive. Le secret est stocké dans une structure JSON avec les clés suivantes :

```
{
    "clientAccount": "service account email",
    "adminAccount": "user account email"",
    "privateKey": "private key"
}
```

- IAM role —Spécifiez à quel `RoleArn` moment vous appelez `CreateDataSource` pour fournir à un IAM rôle les autorisations d'accéder à votre Secrets Manager secret et d'appeler le public requis APIs pour le connecteur Google Drive et Amazon Kendra. Pour plus d'informations, consultez la section IAM Rôles des sources de données Google Drive.

Vous pouvez également ajouter les fonctionnalités optionnelles suivantes :

- Filtres d'inclusion et d'exclusion : Amazon Kendra indexe par défaut tous les documents de Google Drive. Vous pouvez spécifier si vous souhaitez inclure ou exclure certains contenus dans les lecteurs partagés, les comptes utilisateur, les types MIME de documents et les fichiers. Si vous choisissez d'exclure des comptes utilisateurs, aucun des fichiers du My Drive appartenant au compte n'est indexé. Les fichiers partagés avec l'utilisateur sont indexés sauf si le propriétaire du fichier est également exclu.

> **Note**
>
> La plupart des sources de données utilisent des modèles d'expressions régulières, qui sont des modèles d'inclusion ou d'exclusion appelés filtres. Si vous spécifiez un filtre d'inclusion, seul le contenu correspondant au filtre d'inclusion est indexé. Tout document qui ne correspond pas au filtre d'inclusion n'est pas indexé. Si vous spécifiez un filtre d'inclusion et d'exclusion, les documents correspondant au filtre d'exclusion ne sont pas indexés, même s'ils correspondent au filtre d'inclusion.

- Mappages de champs : choisissez de mapper les champs de votre source de données Google Drive à vos champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.

  > **Note**
  >
  > Le champ du corps du document ou l'équivalent du corps du document pour vos documents est requis pour Amazon Kendra effectuer une recherche dans vos documents. Vous devez associer le nom du champ du corps du document dans votre source de données au nom du champ d'index`_document_body`. Tous les autres champs sont facultatifs.

- Filtrage du contexte utilisateur et contrôle d'accèsAmazon Kendra : analyse la liste de contrôle d'accès (ACL) de vos documents, si vous disposez d'une ACL pour vos documents. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.

En savoir plus

Pour en savoir plus sur l'intégration Amazon Kendra à votre source de données Google Drive, consultez :

- Commencer à utiliser le connecteur Amazon Kendra Google Drive

Connecteur Google Drive V2.0

Google Drive est un service de stockage de fichiers basé sur le cloud. Vous pouvez l'utiliser Amazon Kendra pour indexer les documents et les commentaires stockés dans les dossiers Drive partagés, Mes Drives et Shared with me de votre source de données Google Drive. Vous pouvez indexer les documents Google Workspace, ainsi que les documents répertoriés dans la section Types de documentation. Vous pouvez également utiliser des filtres d'inclusion et d'exclusion pour indexer le contenu par nom de fichier, type de fichier et chemin de fichier.

> **Note**
>
> Connecteur Google Drive V1.0/ L' DriveConfiguration API Google a pris fin en 2023. Nous vous recommandons de migrer vers ou d'utiliser le connecteur Google Drive TemplateConfiguration V2.0/API.

Pour résoudre les problèmes liés à votre connecteur de source de données Amazon Kendra Google Drive, consultezDépannage des sources de données.

Rubriques

- Fonctionnalités prises en charge
- Prérequis
- Instructions de connexion
- Remarques

Fonctionnalités prises en charge

- Mappages de champs
- Contrôle d'accès utilisateur
- Filtres d'inclusion/exclusion
- Synchronisation complète et incrémentielle du contenu
- Cloud privé virtuel (VPC)

Prérequis

Avant de pouvoir Amazon Kendra indexer votre source de données Google Drive, apportez ces modifications à votre Google Drive et à vos AWS comptes.

Dans Google Drive, assurez-vous d'avoir :

- Vous avez obtenu l'accès par un rôle de super administrateur ou vous êtes un utilisateur doté de privilèges administratifs. Vous n'avez pas besoin d'un rôle de super administrateur si l'accès vous a été accordé par un rôle de super administrateur.
- Identifiants de connexion au compte de service Google Drive configurés contenant l'adresse e-mail de votre compte administrateur, l'adresse e-mail du client (adresse e-mail du compte de service) et

votre clé privée. Consultez la documentation de Google Cloud sur la création et la suppression de clés de compte de service.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

- Vous avez créé un compte de service Google Cloud (un compte doté de l'autorité déléguée pour assumer une identité d'utilisateur) avec l'activation de l'option Activer la délégation à l'échelle du domaine G Suite à des fins server-to-server d'authentification, puis généré une clé privée JSON à l'aide du compte.

> **Note**
>
> La clé privée doit être générée après la création du compte de service.

- Ajout de l'API du SDK d'administration et de l'API Google Drive dans votre compte utilisateur.
- Facultatif : J'ai configuré les informations de connexion Google Drive OAuth 2.0 contenant l'identifiant du client, le secret du client et le jeton d'actualisation en tant qu'informations de connexion pour un utilisateur spécifique. Vous en avez besoin pour analyser les données de chaque compte. Consultez la documentation Google sur l'utilisation de la OAuth version 2.0 pour y accéder APIs.
- Vous avez ajouté (ou demandé à un utilisateur doté d'un rôle de super administrateur d'ajouter) les OAuth étendues suivantes à votre compte de service à l'aide d'un rôle de super administrateur. Ces étendues d'API sont nécessaires pour analyser tous les documents et les informations de contrôle d'accès (ACL) pour tous les utilisateurs d'un domaine Google Workspace :
  - https://www.googleapis.com/auth/Drive.readOnly : affichez et téléchargez tous vos fichiers Google Drive
  - https://www.googleapis.com/auth/drive.metadata.readonly : affiche les métadonnées des fichiers de votre Google Drive

- https://www.googleapis.com/auth/Admin.Directory.Group.ReadOnly : possibilité de récupérer uniquement les informations relatives au groupe, à l'alias du groupe et aux membres. Cela est nécessaire pour l' Amazon Kendra Identity Crawler.
- https://www.googleapis.com/auth/Admin.Directory.User.ReadOnly : possibilité de récupérer uniquement les utilisateurs ou les alias d'utilisateurs. Cela est nécessaire pour répertorier les utilisateurs dans Amazon Kendra Identity Crawler et pour le ACLs paramétrer.
- https://www.googleapis.com/auth/Plateforme cloud : possibilité de générer un jeton d'accès pour récupérer le contenu de fichiers Google Drive volumineux.
- https://www.googleapis.com/auth/forms.body.readonly : possibilité de récupérer des données depuis Google Forms.

  Pour prendre en charge l'API Forms, ajoutez le champ d'application supplémentaire suivant :

  - https://www.googleapis.com/auth/forms.body.readonly

- Il est vérifié que chaque document est unique dans Google Drive et dans les autres sources de données que vous prévoyez d'utiliser pour le même index. Chaque source de données que vous souhaitez utiliser pour un index ne doit pas contenir le même document dans toutes les sources de données. IDs Les documents sont globaux par rapport à un index et doivent être uniques par index.

Dans votre Compte AWS, assurez-vous d'avoir :

- Création d'un Amazon Kendra index et, si vous utilisez l'API, notez l'ID de l'index.
- Vous avez créé un IAM rôle pour votre source de données et, si vous utilisez l'API, notez l'ARN du IAM rôle.

> **Note**
>
> Si vous modifiez votre type d'authentification et vos informations d'identification, vous devez mettre à jour votre IAM rôle pour accéder au bon identifiant AWS Secrets Manager secret.

- Stockez vos informations d'authentification Google Drive dans un AWS Secrets Manager secret et, si vous utilisez l'API, notez l'ARN du secret.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

Si vous n'avez pas de IAM rôle ou de secret existant, vous pouvez utiliser la console pour créer un nouveau IAM rôle et un nouveau Secrets Manager secret lorsque vous connectez votre source de données Google Drive à Amazon Kendra. Si vous utilisez l'API, vous devez fournir l'ARN d'un IAM rôle et d'un Secrets Manager secret existants, ainsi qu'un identifiant d'index.

Instructions de connexion

Pour vous connecter Amazon Kendra à votre source de données Google Drive, vous devez fournir les informations nécessaires sur votre source de données Google Drive afin de Amazon Kendra pouvoir accéder à vos données. Si vous n'avez pas encore configuré Google Drive pour Amazon Kendra voirPrérequis.

Console

Pour vous connecter Amazon Kendra à Google Drive

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la.
2. Dans le volet de navigation de gauche, choisissez Index, puis choisissez l'index que vous souhaitez utiliser dans la liste des index.

   > **Note**
   >
   > Vous pouvez choisir de configurer ou de modifier vos paramètres de contrôle d'accès utilisateur dans les paramètres de l'index.

3. Sur la page de démarrage, choisissez Ajouter une source de données.
4. Sur la page Ajouter une source de données, choisissez le connecteur Google Drive, puis sélectionnez Ajouter un connecteur. Si vous utilisez la version 2 (le cas échéant), choisissez le connecteur Google Drive avec la balise « V2.0 ».

5. Sur la page Spécifier les détails de la source de données, entrez les informations suivantes :

   a. Dans Nom et description, pour Nom de la source de données : entrez le nom de votre source de données. Vous pouvez inclure des tirets, mais pas des espaces.

   b. (Facultatif) Description : entrez une description facultative pour votre source de données.

   c. Dans la langue par défaut : choisissez une langue pour filtrer vos documents pour l'index. Sauf indication contraire, la langue par défaut est l'anglais. La langue spécifiée dans les métadonnées du document remplace la langue sélectionnée.

   d. Dans Balises, pour Ajouter une nouvelle balise : incluez des balises facultatives pour rechercher et filtrer vos ressources ou suivre vos AWS coûts.

   e. Choisissez Suivant.

6. Sur la page Définir l'accès et la sécurité, entrez les informations suivantes :

   a. Autorisation : activez ou désactivez les informations de la liste de contrôle d'accès (ACL) pour vos documents, si vous disposez d'une ACL et que vous souhaitez l'utiliser pour le contrôle d'accès. L'ACL indique les documents auxquels les utilisateurs et les groupes peuvent accéder. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.

   b. Pour l'authentification : choisissez entre un compte de service Google et une authentification OAuth 2.0 en fonction de votre cas d'utilisation.

   c. AWS Secrets Manager secret —Choisissez un secret existant ou créez-en un nouveau Secrets Manager pour stocker vos informations d'authentification Google Drive. Si vous choisissez de créer un nouveau secret, une fenêtre AWS Secrets Manager secrète s'ouvre.

      i. Si vous avez choisi un compte de service Google, saisissez le nom de votre code secret, l'identifiant e-mail de l'utilisateur administrateur ou « utilisateur du compte de service » dans la configuration de votre compte de service (e-mail d'administrateur), l'identifiant e-mail du compte de service (e-mail du client) et la clé privée que vous avez créée dans votre compte de service.

         Enregistrez et ajoutez votre secret

      ii. Si vous avez choisi l'authentification OAuth 2.0, entrez un nom pour votre code secret, votre identifiant client, votre secret client et le jeton d'actualisation que vous avez créé dans votre OAuth compte. L'identifiant de messagerie de l'utilisateur

(utilisateur dont les détails de connexion sont configurés) sera défini comme ACL. Le connecteur ne définit pas les autres informations principales de l'utilisateur/du groupe en tant qu'ACL en raison des limites de l'API.

Enregistrez et ajoutez votre secret.

d. Virtual Private Cloud (VPC) —Vous pouvez choisir d'utiliser un VPC. Dans ce cas, vous devez ajouter des sous-réseaux et des groupes de sécurité VPC.

e. (Pour les utilisateurs authentifiés par un compte de service Google uniquement)

Identity Crawler : spécifiez s'il faut activer l'explorateur Amazon Kendra d'identité. Le robot d'exploration d'identité utilise les informations de la liste de contrôle d'accès (ACL) de vos documents pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Si vous disposez d'une ACL pour vos documents et que vous choisissez de l'utiliser, vous pouvez également choisir d'activer le robot d'exploration Amazon Kendra d'identité pour configurer le filtrage des résultats de recherche par contexte utilisateur. Sinon, si le robot d'identification est désactivé, tous les documents peuvent être consultés publiquement. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'exploration d'identité est désactivé, vous pouvez également utiliser l'PutPrincipalMappingAPI pour télécharger les informations d'accès des utilisateurs et des groupes afin de filtrer le contexte utilisateur.

f. IAM rôle —Choisissez un IAM rôle existant ou créez-en un nouveau IAM pour accéder aux informations d'identification de votre référentiel et indexer le contenu.

> **Note**
>
> IAM les rôles utilisés pour les index ne peuvent pas être utilisés pour les sources de données. Si vous ne savez pas si un rôle existant est utilisé pour un index ou une FAQ, choisissez Créer un nouveau rôle pour éviter les erreurs.

g. Choisissez Suivant.

7. Sur la page Configurer les paramètres de synchronisation, entrez les informations suivantes :

a. Synchroniser le contenu : sélectionnez les options ou le contenu que vous souhaitez analyser. Vous pouvez choisir d'explorer My Drive (dossiers personnels), Shared Drive (dossiers partagés avec vous) ou les deux. Vous pouvez également inclure des commentaires sur les fichiers.

b. Dans Configuration supplémentaire - facultatif Vous pouvez également saisir les informations facultatives suivantes :

   i. Taille maximale des fichiers : définissez la taille maximale MBs des fichiers à analyser.

   ii. E-mail utilisateur : ajoutez les e-mails utilisateur que vous souhaitez inclure ou exclure.

   iii. Lecteurs partagés : ajoutez les noms des lecteurs partagés que vous souhaitez inclure ou exclure.

   iv. Types MIME : ajoutez les types MIME que vous souhaitez inclure ou exclure.

   v. Modèles d'expression régulière d'entité : ajoutez des modèles d'expressions régulières pour inclure ou exclure certaines pièces jointes pour toutes les entités prises en charge. Vous pouvez ajouter jusqu'à 100 motifs.

      Vous pouvez configurer des modèles d'inclusion/exclusion pour le nom de fichier, le type de fichier et le chemin du fichier.

      - Nom du fichier : nom du fichier à inclure ou à exclure. Par exemple, pour indexer un fichier avec un nom`teamroster.txt`, fournissez`teamroster`.
      - Type de fichier : type de fichier à inclure ou à exclure. Par exemple, .pdf .txt .docx.
      - Chemin du fichier : chemin du fichier à inclure ou à exclure. Par exemple, pour indexer des fichiers uniquement dans le dossier `Products list` d'un lecteur, fournissez`/Products list`.

c. Mode de synchronisation : choisissez la manière dont vous souhaitez mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation.

   - Synchronisation complète : Indexez tout le contenu fraîchement, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
   - Nouvelle synchronisation modifiée : Indexez uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour

suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

- Synchronisation nouvelle, modifiée ou supprimée : Indexez uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

> **Important**
>
> L'API Google Drive ne permet pas de récupérer les commentaires d'un fichier définitivement supprimé. Les commentaires des fichiers supprimés sont récupérables. Lorsqu'un fichier est détruit, le connecteur supprime les commentaires de l' Amazon Kendra index.

d. Dans Calendrier d'exécution de la synchronisation, pour Fréquence, choisissez la fréquence à laquelle vous souhaitez synchroniser le contenu de votre source de données et mettre à jour votre index.

e. Dans Synchroniser l'historique des exécutions, choisissez de stocker les rapports générés automatiquement dans et Amazon S3 lors de la synchronisation de votre source de données. Cela est utile pour suivre les problèmes lors de la synchronisation de votre source de données.

f. Choisissez Suivant.

8. Sur la page Définir les mappages de champs, entrez les informations suivantes :

a. Pour les fichiers : sélectionnez parmi les champs de source de données par défaut Amazon Kendra générés que vous souhaitez mapper à votre index.

> **Note**
>
> L'API Google Drive ne permet pas de créer des champs personnalisés. Le mappage de champs personnalisé n'est pas disponible pour le connecteur Google Drive.

b. Choisissez Suivant.

9. Sur la page Réviser et créer, vérifiez que les informations saisies sont correctes, puis sélectionnez Ajouter une source de données. Vous pouvez également choisir de modifier vos informations à partir de cette page. Votre source de données apparaîtra sur la page Sources de données une fois que la source de données aura été ajoutée avec succès.

API

Pour vous connecter Amazon Kendra à Google Drive

Vous devez spécifier un JSON du schéma de source de données à l'aide de l'TemplateConfigurationAPI. Vous devez fournir les informations suivantes :

- Source de données —Spécifiez le type de source de données comme `GOOGLEDRIVEV2` lorsque vous utilisez TemplateConfigurationSchéma JSON. Spécifiez également la source de données comme `TEMPLATE` lorsque vous appelez le CreateDataSourceAPI.
- Type d'authentification : spécifiez si vous souhaitez utiliser l'authentification du compte de service ou l'authentification OAuth 2.0.
- Mode de synchronisation : spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation. Vous pouvez choisir entre :
  - `FORCED_FULL_CRAWL`pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
  - `FULL_CRAWL`pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
  - `CHANGE_LOG`pour indexer uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

> **⚠ Important**
>
> L'API Google Drive ne permet pas de récupérer les commentaires d'un fichier définitivement supprimé. Les commentaires des fichiers supprimés sont récupérables. Lorsqu'un fichier est détruit, le connecteur supprime les commentaires de l' Amazon Kendra index.

- Nom de ressource Amazon (ARN) secret : indiquez le nom de ressource Amazon (ARN) d'un Secrets Manager secret contenant les informations d'authentification que vous avez créées dans votre compte Google Drive. Si vous utilisez l'authentification par compte de service Google, le secret est stocké dans une structure JSON avec les clés suivantes :

```
{
    "clientEmail": "user account email",
    "adminAccountEmail": "service account email",
    "privateKey": "private key"
}
```

  Si vous utilisez l'authentification OAuth 2.0, le secret est stocké dans une structure JSON avec les clés suivantes :

```
{
    "clientID": "OAuth client ID",
    "clientSecret": "client secret",
    "refreshToken": "refresh token"
}
```

- IAM role —Spécifiez à quel `RoleArn` moment vous appelez `CreateDataSource` pour fournir à un IAM rôle les autorisations d'accéder à votre Secrets Manager secret et d'appeler le public requis APIs pour le connecteur Google Drive et Amazon Kendra. Pour plus d'informations, consultez la section IAM Rôles des sources de données Google Drive.

Vous pouvez également ajouter les fonctionnalités optionnelles suivantes :

- Virtual Private Cloud (VPC) **`VpcConfiguration`** —Spécifiez le moment de votre appel. `CreateDataSource` Pour de plus amples informations, veuillez consulter Configuration Amazon Kendra pour utiliser un Amazon VPC.

- Mes lecteurs, lecteurs partagés, commentaires : vous pouvez indiquer si vous souhaitez analyser ces types de contenu.
- Filtres d'inclusion et d'exclusion : vous pouvez spécifier si vous souhaitez inclure ou exclure certains comptes utilisateurs, lecteurs partagés et types MIME.

> **Note**
>
> La plupart des sources de données utilisent des modèles d'expressions régulières, qui sont des modèles d'inclusion ou d'exclusion appelés filtres. Si vous spécifiez un filtre d'inclusion, seul le contenu correspondant au filtre d'inclusion est indexé. Tout document qui ne correspond pas au filtre d'inclusion n'est pas indexé. Si vous spécifiez un filtre d'inclusion et d'exclusion, les documents correspondant au filtre d'exclusion ne sont pas indexés, même s'ils correspondent au filtre d'inclusion.

- Liste de contrôle d'accès (ACL) : indiquez si vous souhaitez analyser les informations ACL de vos documents, si vous disposez d'une ACL et souhaitez l'utiliser pour le contrôle d'accès. L'ACL indique les documents auxquels les utilisateurs et les groupes peuvent accéder. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.
- Identity Crawler : spécifiez s'il faut activer l'explorateur Amazon Kendra d'identité. Le robot d'exploration d'identité utilise les informations de la liste de contrôle d'accès (ACL) de vos documents pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Si vous disposez d'une ACL pour vos documents et que vous choisissez de l'utiliser, vous pouvez également choisir d'activer le robot d'exploration Amazon Kendra d'identité pour configurer le filtrage des résultats de recherche par contexte utilisateur. Sinon, si le robot d'identification est désactivé, tous les documents peuvent être consultés publiquement. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'exploration d'identité est désactivé, vous pouvez également utiliser l'PutPrincipalMappingAPI pour télécharger les informations d'accès des utilisateurs et des groupes afin de filtrer le contexte utilisateur.
- Mappages de champs : choisissez de mapper les champs de votre source de données Google Drive à vos champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.

> **Note**
>
> Le champ du corps du document ou l'équivalent du corps du document pour vos documents est requis pour Amazon Kendra effectuer une recherche dans vos documents. Vous devez associer le nom du champ du corps du document dans votre source de données au nom du champ d'index`_document_body`. Tous les autres champs sont facultatifs.

Pour obtenir la liste des autres clés JSON importantes à configurer, consultez le schéma du modèle Google Drive.

Remarques

- Le mappage des champs personnalisés n'est pas disponible pour le connecteur Google Drive car l'interface utilisateur de Google Drive ne permet pas de créer des champs personnalisés.
- L'API Google Drive ne permet pas de récupérer les commentaires d'un fichier définitivement supprimé. Les commentaires sont toutefois récupérables pour les fichiers mis à la poubelle. Lorsqu'un fichier est détruit, le Amazon Kendra connecteur supprime les commentaires de l' Amazon Kendra index.
- L'API Google Drive ne renvoie pas les commentaires présents dans un fichier .docx.
- Si l'autorisation est accordée à un particulier Google document (document, feuille de calcul, diapositive, etc.) est défini sur Accès général : si vous possédez le lien ou si vous partagez le domaine de votre entreprise, le document ne sera pas visible pour les utilisateurs de la recherche Amazon Kendra tant que l'utilisateur à l'origine de la requête n'aura pas accédé au document.

IBM DB2

> **Note**
>
> IBM DB2Le connecteur reste entièrement pris en charge pour les clients existants jusqu'au 31 mai 2026. Bien que ce connecteur ne soit plus disponible pour les nouveaux utilisateurs, les utilisateurs actuels peuvent continuer à l'utiliser sans interruption. Nous faisons constamment évoluer notre portefeuille de connecteurs afin de proposer des solutions plus évolutives et personnalisables. Pour les intégrations futures, nous vous recommandons

d'explorer le framework de connecteurs personnalisés Amazon Kendra [1], conçu pour prendre en charge un plus large éventail de cas d'utilisation en entreprise avec une flexibilité accrue.

IBM DB2est un système de gestion de base de données relationnelle développé parIBM. Si vous êtes un IBM DB2 utilisateur, vous pouvez l'utiliser Amazon Kendra pour indexer votre source de IBM DB2 données. Le connecteur de source Amazon Kendra IBM DB2 de données prend en charge la version DB2 11.5.7.

Vous pouvez vous connecter Amazon Kendra à votre source de IBM DB2 données à l'aide de la Amazon Kendra console et de l'TemplateConfigurationAPI.

Pour résoudre les problèmes liés à votre connecteur de source de Amazon Kendra IBM DB2 données, consultezDépannage des sources de données.

Rubriques

- Fonctionnalités prises en charge
- Prérequis
- Instructions de connexion
- Remarques

Fonctionnalités prises en charge

- Mappages de champs
- Filtrage du contexte utilisateur
- Filtres d'inclusion/exclusion
- Synchronisation complète et incrémentielle du contenu
- Cloud privé virtuel (VPC)

Prérequis

Avant de pouvoir Amazon Kendra indexer votre source de IBM DB2 données, apportez ces modifications à vos AWS comptes IBM DB2 and.

DansIBM DB2, assurez-vous d'avoir :

- Notez le nom d'utilisateur et le mot de passe de votre base de données.

> **Important**
>
> Il est recommandé de fournir des informations d'identification Amazon Kendra de base de données en lecture seule.

- Vous avez copié l'URL, le port et l'instance de votre hôte de base de données.
- Il est vérifié que chaque document est unique dans IBM DB2 et entre les autres sources de données que vous prévoyez d'utiliser pour le même index. Chaque source de données que vous souhaitez utiliser pour un index ne doit pas contenir le même document dans toutes les sources de données. IDs Les documents sont globaux par rapport à un index et doivent être uniques par index.

Dans votre Compte AWS, assurez-vous d'avoir :

- Création d'un Amazon Kendra index et, si vous utilisez l'API, notez l'ID de l'index.
- Vous avez créé un IAM rôle pour votre source de données et, si vous utilisez l'API, notez l'ARN du IAM rôle.

> **Note**
>
> Si vous modifiez votre type d'authentification et vos informations d'identification, vous devez mettre à jour votre IAM rôle pour accéder au bon identifiant AWS Secrets Manager secret.

- Stockez vos informations d'IBM DB2authentification dans un AWS Secrets Manager secret et, si vous utilisez l'API, notez l'ARN du secret.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

Si vous n'avez pas de IAM rôle ou de secret existant, vous pouvez utiliser la console pour créer un nouveau IAM rôle et un nouveau Secrets Manager secret lorsque vous connectez votre source de IBM DB2 données à Amazon Kendra. Si vous utilisez l'API, vous devez fournir l'ARN d'un IAM rôle et d'un Secrets Manager secret existants, ainsi qu'un identifiant d'index.

Instructions de connexion

Pour vous connecter Amazon Kendra à votre source de IBM DB2 données, vous devez fournir les détails de vos IBM DB2 informations d'identification afin de Amazon Kendra pouvoir accéder à vos données. Si vous n'avez pas encore configuré IBM DB2 pour Amazon Kendra voirPrérequis.

Console

Pour vous connecter Amazon Kendra à IBM DB2

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la.
2. Dans le volet de navigation de gauche, choisissez Index, puis choisissez l'index que vous souhaitez utiliser dans la liste des index.

   > **Note**
   >
   > Vous pouvez choisir de configurer ou de modifier vos paramètres de contrôle d'accès utilisateur dans les paramètres de l'index.

3. Sur la page de démarrage, choisissez Ajouter une source de données.
4. Sur la page Ajouter une source de données, choisissez IBM DB2connecteur, puis sélectionnez Ajouter un connecteur. Si vous utilisez la version 2 (le cas échéant), choisissez le IBM DB2connecteur avec la balise « V2.0 ».
5. Sur la page Spécifier les détails de la source de données, entrez les informations suivantes :

   a. Dans Nom et description, pour Nom de la source de données : entrez le nom de votre source de données. Vous pouvez inclure des traits d'union, mais pas des espaces.
   b. (Facultatif) Description : entrez une description facultative pour votre source de données.
   c. Dans la langue par défaut : choisissez une langue pour filtrer vos documents pour l'index. Sauf indication contraire, la langue par défaut est l'anglais. La langue spécifiée dans les métadonnées du document remplace la langue sélectionnée.
   d. Dans Balises, pour Ajouter une nouvelle balise : incluez des balises facultatives pour rechercher et filtrer vos ressources ou suivre vos AWS coûts.

e. Choisissez Suivant.

6. Sur la page Définir l'accès et la sécurité, entrez les informations suivantes :

   a. Dans Source, entrez les informations suivantes :

   b. Hôte : entrez le nom d'hôte de la base de données.

   c. Port — Entrez le port de base de données.

   d. Instance — Entrez l'instance de base de données.

   e. Activer l'emplacement du certificat SSL : choisissez d'entrer le Amazon S3 chemin d'accès à votre fichier de certificat SSL.

   f. Dans Authentification, entrez les informations suivantes :

      - AWS Secrets Manager secret —Choisissez un secret existant ou créez-en un nouveau Secrets Manager pour stocker vos informations IBM DB2 d'authentification. Si vous choisissez de créer un nouveau secret, une fenêtre AWS Secrets Manager secrète s'ouvre.

        A. Entrez les informations suivantes dans la fenêtre Créer un AWS Secrets Manager secret :

           I. Nom secret : le nom de votre secret. Le préfixe « AmazonKendra - IBM DB2 - » est automatiquement ajouté à votre nom secret.

           II. Pour le nom d'utilisateur et le mot de passe de la base de données, entrez les valeurs d'identification d'authentification que vous avez copiées depuis votre base de données.

        B. Choisissez Enregistrer.

   g. Virtual Private Cloud (VPC) —Vous pouvez choisir d'utiliser un VPC. Dans ce cas, vous devez ajouter des sous-réseaux et des groupes de sécurité VPC.

   h. IAM rôle —Choisissez un IAM rôle existant ou créez-en un nouveau IAM pour accéder aux informations d'identification de votre référentiel et indexer le contenu.

> **Note**
>
> IAM les rôles utilisés pour les index ne peuvent pas être utilisés pour les sources de données. Si vous ne savez pas si un rôle existant est utilisé pour un index ou une FAQ, choisissez Créer un nouveau rôle pour éviter les erreurs.

i. Choisissez Suivant.

7. Sur la page Configurer les paramètres de synchronisation, entrez les informations suivantes :

   a. Dans le champ d'application de la synchronisation, choisissez l'une des options suivantes :

      - Requête SQL —Entrez des instructions de requête SQL telles que les opérations SELECT et JOIN. La taille des requêtes SQL doit être inférieure à 32 Ko. Amazon Kendra analysera tout le contenu de la base de données correspondant à votre requête.
      - Colonne clé primaire : indiquez la clé primaire pour la table de base de données. Cela identifie une table au sein de votre base de données.
      - Colonne de titre : indiquez le nom de la colonne de titre du document dans votre table de base de données.
      - Colonne du corps : indiquez le nom de la colonne du corps du document dans votre table de base de données.

   b. Dans Configuration supplémentaire — facultatif, choisissez l'une des options suivantes pour synchroniser un contenu spécifique au lieu de synchroniser tous les fichiers :

      - Colonnes détectant les modifications : entrez le nom des colonnes qui Amazon Kendra seront utilisées pour détecter les modifications de contenu. Amazon Kendra réindexera le contenu en cas de modification de l'une de ces colonnes.
      - IDsColonne utilisateur : entrez le nom de la colonne contenant l'utilisateur devant être autorisé IDs à accéder au contenu.
      - Colonne Groupes : entrez le nom de la colonne qui contient les groupes autorisés à accéder au contenu.
      - URLsColonne source —Entrez le nom de la colonne contenant la source URLs à indexer.
      - Colonne horodatage : entrez le nom de la colonne contenant les horodatages. Amazon Kendra utilise les informations d'horodatage pour détecter les modifications apportées à votre contenu et synchroniser uniquement le contenu modifié.
      - Colonne fuseaux horaires : entrez le nom de la colonne contenant les fuseaux horaires du contenu à analyser.
      - Format d'horodatage : entrez le nom de la colonne contenant les formats d'horodatage à utiliser pour détecter les modifications de contenu et resynchroniser votre contenu.

c. Mode de synchronisation : choisissez la manière dont vous souhaitez mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation.

   - Synchronisation complète : Indexez tout le contenu fraîchement, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
   - Nouvelle synchronisation modifiée : Indexez uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
   - Synchronisation nouvelle, modifiée ou supprimée : Indexez uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

d. Dans Synchroniser le calendrier d'exécution, pour Fréquence : à quelle fréquence Amazon Kendra se synchronisera avec votre source de données.

e. Choisissez Suivant.

8. Sur la page Définir les mappages de champs, entrez les informations suivantes :

   a. Sélectionnez parmi les champs de source de données par défaut générés (Document IDs, Titres des documents et Source URLs) que vous souhaitez Amazon Kendra mapper à l'index.

   b. Ajouter un champ —Pour ajouter des champs de source de données personnalisés afin de créer un nom de champ d'index à mapper et le type de données du champ.

   c. Choisissez Suivant.

9. Sur la page Réviser et créer, vérifiez que les informations saisies sont correctes, puis sélectionnez Ajouter une source de données. Vous pouvez également choisir de modifier vos informations à partir de cette page. Votre source de données apparaîtra sur la page Sources de données une fois que la source de données aura été ajoutée avec succès.

API

Pour vous connecter Amazon Kendra à IBM DB2

Vous devez spécifier les éléments suivants à l'aide de l'TemplateConfigurationAPI :

- Source de données —Spécifiez le type de source de données tel que `JDBC` lorsque vous utilisez le schéma TemplateConfigurationJSON. Spécifiez également la source de données `TEMPLATE` lorsque vous appelez l'CreateDataSourceAPI.
- Type de base de données —Vous devez spécifier le type de base de données sous `db2` la forme.
- Requête SQL : spécifiez des instructions de requête SQL telles que les opérations SELECT et JOIN. La taille des requêtes SQL doit être inférieure à 32 Ko. Amazon Kendra analysera tout le contenu de la base de données correspondant à votre requête.
- Mode de synchronisation : spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation. Vous pouvez choisir entre :
  - `FORCED_FULL_CRAWL`pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
  - `FULL_CRAWL`pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
  - `CHANGE_LOG`pour indexer uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
- Nom de ressource Amazon (ARN) secret : indiquez le nom de ressource Amazon (ARN) d'un Secrets Manager secret contenant les informations d'authentification que vous avez créées dans votre IBM DB2 compte. Le secret est stocké dans une structure JSON avec les clés suivantes :

```
{
    "user name": "database user name",
```

```
    "password": "password"
}
```

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

- IAM role —Spécifiez à quel `RoleArn` moment vous appelez `CreateDataSource` pour fournir à un IAM rôle les autorisations d'accéder à votre Secrets Manager secret et d'appeler le public requis APIs pour le IBM DB2 connecteur et Amazon Kendra. Pour plus d'informations, consultez la section IAM Rôles des sources de IBM DB2 données.

Vous pouvez également ajouter les fonctionnalités optionnelles suivantes :

- Virtual Private Cloud (VPC) **`VpcConfiguration`** —Spécifiez le moment de votre appel. `CreateDataSource` Pour de plus amples informations, veuillez consulter Configuration Amazon Kendra pour utiliser un Amazon VPC.
- Filtres d'inclusion et d'exclusion : vous pouvez spécifier si vous souhaitez inclure un contenu spécifique en utilisant l'utilisateur IDs, les groupes, la source URLs, les horodatages et les fuseaux horaires.
- Filtrage du contexte utilisateur et contrôle d'accèsAmazon Kendra : analyse la liste de contrôle d'accès (ACL) de vos documents, si vous disposez d'une ACL pour vos documents. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.
- Mappages de champs : choisissez de mapper les champs de votre source de IBM DB2 données à vos champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.

> **Note**
>
> Le champ du corps du document ou l'équivalent du corps du document pour vos documents est requis pour Amazon Kendra effectuer une recherche dans vos

documents. Vous devez associer le nom du champ du corps du document dans votre source de données au nom du champ d'index`_document_body`. Tous les autres champs sont facultatifs.

Pour obtenir la liste des autres clés JSON importantes à configurer, consultezSchéma DB2 du modèle IBM.

Remarques

- Les lignes de base de données supprimées ne seront pas suivies lors de la Amazon Kendra vérification du contenu mis à jour.
- La taille des noms et des valeurs des champs d'une ligne de votre base de données ne peut pas dépasser 400 Ko.
- Si la source de données de votre base de données contient une grande quantité de données et que vous ne souhaitez pas Amazon Kendra indexer tout le contenu de votre base de données après la première synchronisation, vous pouvez choisir de ne synchroniser que les documents nouveaux, modifiés ou supprimés.
- Il est recommandé de fournir des informations d'identification Amazon Kendra de base de données en lecture seule.
- Il est recommandé d'éviter d'ajouter des tableaux contenant des données sensibles ou des informations personnelles identifiables (PII).

Jira

Jira est un outil de gestion de projet pour le développement de logiciels, la gestion de produits et le suivi des bogues. Vous pouvez l'utiliser Amazon Kendra pour indexer vos projets, problèmes, commentaires, pièces jointes, journaux de travail et statuts Jira.

Amazon Kendra ne prend actuellement en charge que Jira Cloud.

Vous pouvez vous connecter Amazon Kendra à votre source de données Jira à l'aide de la Amazon Kendra console ou de l'JiraConfiguration API. Pour obtenir la liste des fonctionnalités prises en charge par chacune d'entre elles, consultezFonctionnalités prises en charge.

Pour résoudre les problèmes liés à votre connecteur de source de données Amazon Kendra Jira, consultezDépannage des sources de données.

Rubriques

- Fonctionnalités prises en charge
- Prérequis
- Instructions de connexion
- En savoir plus

Fonctionnalités prises en charge

Amazon Kendra Le connecteur de source de données Jira prend en charge les fonctionnalités suivantes :

- Mappages de champs
- Contrôle d'accès des utilisateurs
- Filtres d'inclusion/exclusion
- Synchronisation complète et incrémentielle du contenu
- Cloud privé virtuel (VPC)

Prérequis

Avant de pouvoir Amazon Kendra indexer votre source de données Jira, apportez ces modifications à votre Jira et AWS à vos comptes.

Dans Jira, assurez-vous que vous disposez des éléments suivants :

- Identifiants d'authentification par jeton d'API configurés, qui incluent un identifiant Jira (nom d'utilisateur ou e-mail) et un identifiant Jira (jeton d'API Jira). Consultez la documentation Atlassian sur la gestion des jetons d'API.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

- Vous avez noté l'URL du compte Jira dans les paramètres de votre compte Jira. Par exemple, *https://company.atlassian.net/*.
- Il est vérifié que chaque document est unique dans Jira et dans les autres sources de données que vous prévoyez d'utiliser pour le même index. Chaque source de données que vous souhaitez utiliser pour un index ne doit pas contenir le même document dans toutes les sources de données. IDs Les documents sont globaux par rapport à un index et doivent être uniques par index.

Dans votre Compte AWS, assurez-vous d'avoir :

- Création d'un Amazon Kendra index et, si vous utilisez l'API, notez l'ID de l'index.
- Vous avez créé un IAM rôle pour votre source de données et, si vous utilisez l'API, notez l'ARN du IAM rôle.

> **Note**
>
> Si vous modifiez votre type d'authentification et vos informations d'identification, vous devez mettre à jour votre IAM rôle pour accéder au bon identifiant AWS Secrets Manager secret.

- Stockez vos informations d'authentification Jira dans un AWS Secrets Manager secret et, si vous utilisez l'API, notez l'ARN du secret.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

Si vous n'avez pas de IAM rôle ou de secret existant, vous pouvez utiliser la console pour créer un nouveau IAM rôle et un nouveau Secrets Manager secret lorsque vous connectez votre source de données Jira à Amazon Kendra. Si vous utilisez l'API, vous devez fournir l'ARN d'un IAM rôle et d'un Secrets Manager secret existants, ainsi qu'un identifiant d'index.

Instructions de connexion

Pour vous connecter Amazon Kendra à votre source de données Jira, vous devez fournir les informations nécessaires sur votre source de données Jira afin de Amazon Kendra pouvoir accéder à vos données. Si vous n'avez pas encore configuré Jira pour Amazon Kendra, consultezPrérequis.

Console

Pour vous connecter Amazon Kendra à Jira

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la.
2. Dans le volet de navigation de gauche, choisissez Index, puis choisissez l'index que vous souhaitez utiliser dans la liste des index.

> **Note**
>
> Vous pouvez choisir de configurer ou de modifier vos paramètres de contrôle d'accès utilisateur dans les paramètres de l'index.

3. Sur la page de démarrage, choisissez Ajouter une source de données.
4. Sur la page Ajouter une source de données, choisissez le connecteur Jira, puis sélectionnez Ajouter un connecteur. Si vous utilisez la version 2 (le cas échéant), choisissez le connecteur Jira avec le tag « V2.0 ».
5. Sur la page Spécifier les détails de la source de données, entrez les informations suivantes :
   a. Dans Nom et description, pour Nom de la source de données : entrez le nom de votre source de données. Vous pouvez inclure des traits d'union, mais pas des espaces.
   b. (Facultatif) Description : entrez une description facultative pour votre source de données.
   c. Dans la langue par défaut : choisissez une langue pour filtrer vos documents pour l'index. Sauf indication contraire, la langue par défaut est l'anglais. La langue spécifiée dans les métadonnées du document remplace la langue sélectionnée.
   d. Dans Balises, pour Ajouter une nouvelle balise : incluez des balises facultatives pour rechercher et filtrer vos ressources ou suivre vos AWS coûts.
   e. Choisissez Suivant.
6. Sur la page Définir l'accès et la sécurité, entrez les informations suivantes :

a. URL du compte Jira —Entrez l'URL de votre compte Jira. olpPar exemple : *https://company.atlassian.net/*.

b. Autorisation : activez ou désactivez les informations de la liste de contrôle d'accès (ACL) pour vos documents, si vous disposez d'une ACL et que vous souhaitez l'utiliser pour le contrôle d'accès. L'ACL indique les documents auxquels les utilisateurs et les groupes peuvent accéder. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.

c. AWS Secrets Manager secret —Choisissez un secret existant ou créez-en un nouveau Secrets Manager pour stocker vos informations d'authentification Jira. Si vous choisissez de créer un nouveau secret, une fenêtre AWS Secrets Manager secrète s'ouvre.

   i. Entrez les informations suivantes dans la fenêtre Créer un AWS Secrets Manager secret :

      A. Nom secret : le nom de votre secret. Le préfixe « AmazonKendra -Jira » est automatiquement ajouté à votre nom secret.

      B. Pour Jira ID : entrez le nom d'utilisateur ou l'e-mail Jira.

      C. Pour le mot de passe/le jeton : entrez le jeton d'API Jira configuré dans Jira.

   ii. Enregistrez et ajoutez votre secret.

d. Virtual Private Cloud (VPC) —Vous pouvez choisir d'utiliser un VPC. Dans ce cas, vous devez ajouter des sous-réseaux et des groupes de sécurité VPC.

e. Identity Crawler : spécifiez s'il faut activer l'explorateur Amazon Kendra d'identité. Le robot d'exploration d'identité utilise les informations de la liste de contrôle d'accès (ACL) de vos documents pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Si vous disposez d'une ACL pour vos documents et que vous choisissez de l'utiliser, vous pouvez également choisir d'activer le robot d'exploration Amazon Kendra d'identité pour configurer le filtrage des résultats de recherche par contexte utilisateur. Dans le cas contraire, si Identity Crawler est désactivé, tous les documents peuvent faire l'objet d'une recherche publique. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'exploration d'identité est désactivé, vous pouvez également utiliser l'PutPrincipalMappingAPI pour télécharger les informations d'accès des utilisateurs et des groupes afin de filtrer le contexte utilisateur.

f. IAM rôle —Choisissez un IAM rôle existant ou créez-en un nouveau IAM pour accéder aux informations d'identification de votre référentiel et indexer le contenu.

> **Note**
>
> IAM les rôles utilisés pour les index ne peuvent pas être utilisés pour les sources de données. Si vous ne savez pas si un rôle existant est utilisé pour un index ou une FAQ, choisissez Créer un nouveau rôle pour éviter les erreurs.

g. Choisissez Suivant.

7. Sur la page Configurer les paramètres de synchronisation, entrez les informations suivantes :

a. Sélectionnez les projets Jira à indexer : choisissez d'explorer tous les projets ou des projets spécifiques.

b. Configuration supplémentaire : spécifiez certains statuts et types de problèmes. Choisissez d'explorer les commentaires, les pièces jointes et les journaux de travail. Utilisez des modèles d'expressions régulières pour inclure ou exclure certains contenus.

c. Mode de synchronisation : choisissez la manière dont vous souhaitez mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation.

- Synchronisation complète : Indexez tout le contenu fraîchement, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
- Nouvelle synchronisation modifiée : Indexez uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
- Synchronisation nouvelle, modifiée ou supprimée : Indexez uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre

source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

d. Dans Calendrier d'exécution de la synchronisation, pour Fréquence : choisissez la fréquence à laquelle vous souhaitez synchroniser le contenu de votre source de données et mettre à jour votre index.

e. Choisissez Suivant.

8. Sur la page Définir les mappages de champs, entrez les informations suivantes :

a. Champs de source de données par défaut : sélectionnez parmi les champs de source de données par défaut Amazon Kendra générés que vous souhaitez mapper à votre index.

b. Ajouter un champ —Pour ajouter des champs de source de données personnalisés afin de créer un nom de champ d'index à mapper et le type de données du champ.

c. Choisissez Suivant.

9. Sur la page Réviser et créer, vérifiez que les informations saisies sont correctes, puis sélectionnez Ajouter une source de données. Vous pouvez également choisir de modifier vos informations à partir de cette page. Votre source de données apparaîtra sur la page Sources de données une fois que la source de données aura été ajoutée avec succès.

API

Pour vous connecter Amazon Kendra à Jira

Vous devez spécifier les éléments suivants à l'aide de l'JiraConfiguration API :

- URL de la source de données : spécifiez l'URL de votre compte Jira. Par exemple, *company.atlassian.net*.
- Nom de ressource Amazon (ARN) secret : indiquez le nom de ressource Amazon (ARN) d'un Secrets Manager secret contenant les informations d'authentification de votre compte Jira. Le secret est stocké dans une structure JSON avec les clés suivantes :

```
{
    "jiraId": "Jira user name or email",
    "jiraCredential": "Jira API token"
}
```

- IAM role —Spécifiez à quel `RoleArn` moment vous appelez `CreateDataSource` pour fournir à un IAM rôle les autorisations d'accéder à votre Secrets Manager secret et pour appeler

le public requis APIs pour le connecteur Jira et. Amazon Kendra Pour plus d'informations, consultez la section IAM Rôles des sources de données Jira.

Vous pouvez également ajouter les fonctionnalités optionnelles suivantes :

- Virtual Private Cloud (VPC) —Spécifiez dans le `VpcConfiguration` cadre de la configuration de la source de données. Reportez-vous Amazon Kendra à la section Configuration pour utiliser un VPC.
- Journal des modifications : Amazon Kendra faut-il utiliser le mécanisme du journal des modifications de la source de données Jira pour déterminer si un document doit être mis à jour dans l'index.

  > **Note**
  >
  > Utilisez le journal des modifications si vous ne Amazon Kendra souhaitez pas numériser tous les documents. Si votre journal des modifications est volumineux, la numérisation des documents de la source de données Jira peut prendre Amazon Kendra moins de temps que le traitement du journal des modifications. Si vous synchronisez votre source de données Jira avec votre index pour la première fois, tous les documents sont numérisés.

- Filtres d'inclusion et d'exclusion : vous pouvez indiquer si vous souhaitez inclure ou exclure certains fichiers.

  > **Note**
  >
  > La plupart des sources de données utilisent des modèles d'expressions régulières, qui sont des modèles d'inclusion ou d'exclusion appelés filtres. Si vous spécifiez un filtre d'inclusion, seul le contenu correspondant au filtre d'inclusion est indexé. Tout document qui ne correspond pas au filtre d'inclusion n'est pas indexé. Si vous spécifiez un filtre d'inclusion et d'exclusion, les documents correspondant au filtre d'exclusion ne sont pas indexés, même s'ils correspondent au filtre d'inclusion.

- Commentaires, pièces jointes et journaux de travail : vous pouvez indiquer si vous souhaitez analyser certains commentaires, pièces jointes et journaux de travail relatifs aux problèmes.
- Projets, problèmes, statuts : vous pouvez spécifier si vous souhaitez analyser certains projets IDs, types de problèmes et statuts.

- Filtrage du contexte utilisateur et contrôle d'accèsAmazon Kendra : analyse la liste de contrôle d'accès (ACL) de vos documents, si vous disposez d'une ACL pour vos documents. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.
- Mappages de champs : choisissez de mapper les champs de votre source de données Jira à vos Amazon Kendra champs d'index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.

> **Note**
>
> Le champ du corps du document ou l'équivalent du corps du document pour vos documents est requis pour Amazon Kendra effectuer une recherche dans vos documents. Vous devez associer le nom du champ du corps du document dans votre source de données au nom du champ d'index`_document_body`. Tous les autres champs sont facultatifs.

En savoir plus

Pour en savoir plus sur l'intégration Amazon Kendra à votre source de données Jira, consultez :

- Effectuez des recherches intelligentes dans vos projets Jira avec le connecteur Amazon Kendra Jira Cloud

Microsoft Exchange

Microsoft Exchange est un outil de collaboration d'entreprise pour la messagerie, les réunions et le partage de fichiers. Si vous êtes un utilisateur de Microsoft Exchange, vous pouvez l'utiliser Amazon Kendra pour indexer votre source de données Microsoft Exchange.

Vous pouvez vous connecter Amazon Kendra à votre source de données Microsoft Exchange à l'aide de la Amazon Kendra console et de l'TemplateConfigurationAPI.

Pour résoudre les problèmes liés à votre connecteur de source de données Amazon Kendra Microsoft Exchange, consultezDépannage des sources de données.

Fonctionnalités prises en charge

- Mappages de champs
- Contrôle d'accès des utilisateurs
- Filtres d'inclusion/exclusion
- Synchronisation complète et incrémentielle du contenu
- Cloud privé virtuel (VPC)

Prérequis

Avant de pouvoir Amazon Kendra indexer votre source de données Microsoft Exchange, apportez ces modifications à votre Microsoft Exchange et à vos AWS comptes.

Dans Microsoft Exchange, assurez-vous que vous disposez des éléments suivants :

- Création d'un compte Microsoft Exchange dans Office 365.
- Vous avez noté votre identifiant de client Microsoft 365. Vous pouvez trouver votre ID de locataire dans les propriétés de votre portail Azure Active Directory ou dans votre OAuth application.
- J'ai configuré une OAuth application sur le portail Azure et noté l'ID du client et le secret du client ou les informations d'identification du client. Consultez le didacticiel Microsoft et l'exemple d'application enregistrée pour plus d'informations.

> **Note**
>
> Lorsque vous créez ou enregistrez une application sur le portail Azure, l'ID secret représente la valeur secrète réelle. Vous devez prendre note ou enregistrer la valeur secrète réelle immédiatement lors de la création du secret et de l'application. Vous pouvez accéder à votre secret en sélectionnant le nom de votre application sur le portail Azure, puis en accédant à l'option de menu sur les certificats et les secrets.
> Vous pouvez accéder à votre ID client en sélectionnant le nom de votre application sur le portail Azure, puis en accédant à la page de présentation. L'ID de l'application (client) est l'ID du client.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

- Les autorisations suivantes ont été ajoutées pour l'application de connecteur :

Microsoft Graph	Office 365 Exchange Online
• Mail.Read (Application) • Courrier. ReadBasic (Demande) • Courrier. ReadBasic.Tout (Application) • Calendriers.Read (Application) • Utilisateur.Lisez.Tout (Application) • Contacts.Read (Application) • Remarques.Lisez.Tout (Application) • Directory.Read.Tout (Application) • NOUVELLES. AccessAsUser.Tout (délégué)	accès complet en tant qu'application (Application)

- Il est vérifié que chaque document est unique dans Microsoft Exchange et dans les autres sources de données que vous prévoyez d'utiliser pour le même index. Chaque source de données que vous souhaitez utiliser pour un index ne doit pas contenir le même document dans toutes les sources de données. IDs Les documents sont globaux par rapport à un index et doivent être uniques par index.

Dans votre Compte AWS, assurez-vous d'avoir :

- Création d'un Amazon Kendra index et, si vous utilisez l'API, notez l'ID de l'index.
- Vous avez créé un IAM rôle pour votre source de données et, si vous utilisez l'API, notez l'ARN du IAM rôle.

> **Note**
>
> Si vous modifiez votre type d'authentification et vos informations d'identification, vous devez mettre à jour votre IAM rôle pour accéder au bon identifiant AWS Secrets Manager secret.

- Stockez vos informations d'authentification Microsoft Exchange dans un AWS Secrets Manager secret et, si vous utilisez l'API, notez l'ARN du secret.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

Si vous n'avez pas de IAM rôle ou de secret existant, vous pouvez utiliser la console pour créer un nouveau IAM rôle et un nouveau Secrets Manager secret lorsque vous connectez votre source de données Microsoft Exchange à Amazon Kendra. Si vous utilisez l'API, vous devez fournir l'ARN d'un IAM rôle et d'un Secrets Manager secret existants, ainsi qu'un identifiant d'index.

Instructions de connexion

Pour vous connecter Amazon Kendra à votre source de données Microsoft Exchange, vous devez fournir les informations nécessaires sur votre source de données Microsoft Exchange afin de Amazon Kendra pouvoir accéder à vos données. Si vous n'avez pas encore configuré Microsoft Exchange pour Amazon Kendra, consultezPrérequis.

Console

Pour vous connecter Amazon Kendra à Microsoft Exchange

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la.
2. Dans le volet de navigation de gauche, choisissez Index, puis choisissez l'index que vous souhaitez utiliser dans la liste des index.

> **Note**
>
> Vous pouvez choisir de configurer ou de modifier vos paramètres de contrôle d'accès utilisateur dans les paramètres de l'index.

3. Sur la page de démarrage, choisissez Ajouter une source de données.
4. Sur la page Ajouter une source de données, choisissez le connecteur Microsoft Exchange, puis sélectionnez Ajouter un connecteur. Si vous utilisez la version 2 (le cas échéant), choisissez le connecteur Microsoft Exchange avec la balise « V2.0 ».
5. Sur la page Spécifier les détails de la source de données, entrez les informations suivantes :

   a. Dans Nom et description, pour Nom de la source de données : entrez le nom de votre source de données. Vous pouvez inclure des traits d'union, mais pas des espaces.

   b. (Facultatif) Description : entrez une description facultative pour votre source de données.

   c. Dans la langue par défaut : choisissez une langue pour filtrer vos documents pour l'index. Sauf indication contraire, la langue par défaut est l'anglais. La langue spécifiée dans les métadonnées du document remplace la langue sélectionnée.

   d. Dans Balises, pour Ajouter une nouvelle balise : incluez des balises facultatives pour rechercher et filtrer vos ressources ou suivre vos AWS coûts.

   e. Choisissez Suivant.
6. Sur la page Définir l'accès et la sécurité, entrez les informations suivantes :

   a. ID de locataire —Entrez votre identifiant de locataire Microsoft 365. Vous pouvez trouver votre ID de locataire dans les propriétés de votre portail Azure Active Directory ou dans votre OAuth application.

   b. Autorisation : activez ou désactivez les informations de la liste de contrôle d'accès (ACL) pour vos documents, si vous disposez d'une ACL et que vous souhaitez l'utiliser pour le contrôle d'accès. L'ACL indique les documents auxquels les utilisateurs et les groupes peuvent accéder. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.

   c. AWS Secrets Manager secret —Choisissez un secret existant ou créez-en un nouveau Secrets Manager pour stocker vos informations d'authentification Microsoft Exchange. Si vous choisissez de créer un nouveau secret, une fenêtre AWS Secrets Manager secrète s'ouvre.

i. Entrez les informations suivantes dans la fenêtre Créer un AWS Secrets Manager secret :

A. Nom secret : le nom de votre secret. Le préfixe « AmazonKendra -Microsoft Exchange

B. Pour l'ID client, secret du client : entrez les informations d'authentification configurées dans Microsoft Exchange sur le portail Azure.

ii. Enregistrez et ajoutez votre secret.

d. Virtual Private Cloud (VPC) —Vous pouvez choisir d'utiliser un VPC. Dans ce cas, vous devez ajouter des sous-réseaux et des groupes de sécurité VPC.

e. IAM rôle —Choisissez un IAM rôle existant ou créez-en un nouveau IAM pour accéder aux informations d'identification de votre référentiel et indexer le contenu.

> **Note**
>
> IAM les rôles utilisés pour les index ne peuvent pas être utilisés pour les sources de données. Si vous ne savez pas si un rôle existant est utilisé pour un index ou une FAQ, choisissez Créer un nouveau rôle pour éviter les erreurs.

f. Choisissez Suivant.

7. Sur la page Configurer les paramètres de synchronisation, entrez les informations suivantes :

a. Utilisateur IDs : fournissez les e-mails de l'utilisateur si vous souhaitez filtrer le contenu en fonction de certains e-mails.

b. Configuration supplémentaire —Spécifiez les types de contenu que vous souhaitez analyser.

- Types d'entités : vous pouvez choisir d'explorer le contenu du calendrier ou OneNotes des contacts.
- Analyse du calendrier : entrez les dates de début et de fin pour analyser le contenu entre certaines dates.
- Inclure l'e-mail : saisissez « à », « de » et les lignes d'objet de l'e-mail pour filtrer certains e-mails que vous souhaitez analyser.
- Accès aux dossiers partagés : choisissez d'activer l'analyse de la liste de contrôle d'accès pour contrôler l'accès à votre source de données Microsoft Exchange.

- Regex pour les domaines : ajoutez des modèles d'expressions régulières pour inclure ou exclure certains domaines de messagerie.
- Modèles Regex : ajoutez des modèles d'expressions régulières pour inclure ou exclure certains fichiers.

c. Mode de synchronisation : choisissez la manière dont vous souhaitez mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation.

- Synchronisation complète : Indexez tout le contenu fraîchement, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
- Nouvelle synchronisation modifiée : Indexez uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
- Synchronisation nouvelle, modifiée ou supprimée : Indexez uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

d. Dans Calendrier d'exécution de la synchronisation, pour Fréquence : choisissez la fréquence à laquelle vous souhaitez synchroniser le contenu de votre source de données et mettre à jour votre index.

e. Choisissez Suivant.

8. Sur la page Définir les mappages de champs, entrez les informations suivantes :

a. Champs de source de données par défaut : sélectionnez parmi les champs de source de données par défaut Amazon Kendra générés que vous souhaitez mapper à votre index.

> **Note**
>
> Le connecteur de source de données Amazon Kendra Microsoft Exchange ne prend pas en charge les mappages de champs personnalisés.

b. Choisissez Suivant.

9. Sur la page Réviser et créer, vérifiez que les informations saisies sont correctes, puis sélectionnez Ajouter une source de données. Vous pouvez également choisir de modifier vos informations à partir de cette page. Votre source de données apparaîtra sur la page Sources de données une fois que la source de données aura été ajoutée avec succès.

API

Pour vous connecter Amazon Kendra à Microsoft Exchange

Vous devez spécifier un JSON du schéma de source de données à l'aide de l'TemplateConfigurationAPI. Vous devez fournir les informations suivantes :

- Source de données —Spécifiez le type de source de données tel que `MSEXCHANGE` lorsque vous utilisez le schéma TemplateConfigurationJSON. Spécifiez également la source de données `TEMPLATE` lorsque vous appelez l'CreateDataSourceAPI.
- ID de locataire : vous pouvez trouver votre identifiant de locataire dans les propriétés de votre portail Azure Active Directory ou dans votre OAuth application.
- Mode de synchronisation : spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation. Vous pouvez choisir entre :
  - `FORCED_FULL_CRAWL`pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
  - `FULL_CRAWL`pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

  - `CHANGE_LOG`pour indexer uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
- Nom de ressource Amazon (ARN) secret : indiquez le nom de ressource Amazon (ARN) d'un Secrets Manager secret contenant les informations d'authentification de votre compte Microsoft Exchange. Le secret est stocké dans une structure JSON avec les clés suivantes :

```
{
    "clientId": "client ID",
    "clientSecret": "client secret"
}
```

- IAM role —Spécifiez à quel `RoleArn` moment vous appelez `CreateDataSource` pour fournir à un IAM rôle les autorisations d'accéder à votre Secrets Manager secret et pour appeler le public requis APIs pour le connecteur Microsoft Exchange et Amazon Kendra. Pour plus d'informations, consultez la section IAM Rôles des sources de données Microsoft Exchange.

Vous pouvez également ajouter les fonctionnalités optionnelles suivantes :

- Virtual Private Cloud (VPC) **`VpcConfiguration`** —Spécifiez le moment de votre appel. `CreateDataSource` Pour de plus amples informations, veuillez consulter Configuration Amazon Kendra pour utiliser un Amazon VPC.
- Filtres d'inclusion et d'exclusion : spécifiez s'il faut inclure ou exclure certains contenus.

> **Note**
>
> La plupart des sources de données utilisent des modèles d'expressions régulières, qui sont des modèles d'inclusion ou d'exclusion appelés filtres. Si vous spécifiez un filtre d'inclusion, seul le contenu correspondant au filtre d'inclusion est indexé. Tout document qui ne correspond pas au filtre d'inclusion n'est pas indexé. Si vous spécifiez un filtre d'inclusion et d'exclusion, les documents correspondant au filtre d'exclusion ne sont pas indexés, même s'ils correspondent au filtre d'inclusion.

- Liste de contrôle d'accès (ACL) : indiquez si vous souhaitez analyser les informations ACL de vos documents, si vous disposez d'une ACL et souhaitez l'utiliser pour le contrôle d'accès. L'ACL indique les documents auxquels les utilisateurs et les groupes peuvent accéder. Les

informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.

- Mappages de champs : choisissez de mapper les champs de votre source de données Microsoft Exchange à vos champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.

> **Note**
>
> Le champ du corps du document ou l'équivalent du corps du document pour vos documents est requis pour Amazon Kendra effectuer une recherche dans vos documents. Vous devez associer le nom du champ du corps du document dans votre source de données au nom du champ d'index`_document_body`. Tous les autres champs sont facultatifs.

Pour obtenir la liste des autres clés JSON importantes à configurer, consultez le schéma du modèle Microsoft Exchange.

En savoir plus

Pour en savoir plus sur l'intégration Amazon Kendra à votre source de données Microsoft Exchange, consultez :

- Indexez votre contenu Microsoft Exchange à l'aide du connecteur Exchange pour Amazon Kendra

Remarques

- Lorsque les listes de contrôle d'accès (ACLs) sont activées, l'option « Synchroniser uniquement le contenu nouveau ou modifié » n'est pas disponible en raison des limites de l'API Microsoft Exchange. Nous vous recommandons d'utiliser plutôt les modes « Synchronisation complète » ou « Synchronisation du contenu nouveau, modifié ou supprimé », ou de les désactiver ACLs si vous devez utiliser ce mode de synchronisation.

Microsoft OneDrive

Microsoft OneDrive est un service de stockage basé sur le cloud que vous pouvez utiliser pour stocker, partager et héberger votre contenu. Vous pouvez l'utiliser Amazon Kendra pour indexer votre source de OneDrive données.

Vous pouvez vous connecter Amazon Kendra à votre source de OneDrive données à l'aide de la Amazon Kendra console et de l'OneDriveConfigurationAPI.

Amazon Kendra possède deux versions du OneDrive connecteur. Les fonctionnalités prises en charge par chaque version incluent :

OneDrive Connecteur Microsoft V1.0/API OneDriveConfiguration

- Mappages de champs
- Filtres d'inclusion/exclusion

OneDrive Connecteur Microsoft V2.0/API TemplateConfiguration

- Filtrage du contexte utilisateur
- explorateur d'identité utilisateur
- Filtres d'inclusion/exclusion
- Synchronisation complète et incrémentielle du contenu
- Cloud privé virtuel (VPC)

> **Note**
>
> Support pour le OneDrive connecteur V1.0/ OneDriveConfiguration API devrait prendre fin en juin 2023. Nous vous recommandons d'utiliser le OneDrive connecteur TemplateConfiguration V2.0/API.

Pour résoudre les problèmes liés à votre connecteur de source de Amazon Kendra OneDrive données, consultezDépannage des sources de données.

Rubriques

- OneDrive Connecteur Microsoft V1.0

- OneDrive Connecteur Microsoft V2.0
- En savoir plus
- Remarques

OneDrive Connecteur Microsoft V1.0

Microsoft OneDrive est un service de stockage basé sur le cloud que vous pouvez utiliser pour stocker, partager et héberger votre contenu. Vous pouvez l'utiliser Amazon Kendra pour indexer votre source OneDrive de données Microsoft.

> **Note**
>
> Support pour le OneDrive connecteur V1.0/ l' OneDrive API Microsoft devrait prendre fin en juin 2023. Nous vous recommandons d'utiliser le OneDrive connecteur TemplateConfiguration V2.0/API.

Pour résoudre les problèmes liés à votre connecteur de source de Amazon Kendra OneDrive données, consultezDépannage des sources de données.

Rubriques

- Fonctionnalités prises en charge
- Prérequis
- Instructions de connexion

Fonctionnalités prises en charge

- Mappages de champs
- Filtres d'inclusion/exclusion

Prérequis

Avant de pouvoir Amazon Kendra indexer votre source de OneDrive données, apportez ces modifications à vos AWS comptes OneDrive and.

Dans votre Azure Active Directory (AD), assurez-vous que vous disposez des éléments suivants :

- Création d'une application Azure Active Directory (AD).
- A utilisé l'ID de l'application AD pour enregistrer une clé secrète pour l'application sur le site AD. La clé secrète doit contenir l'ID de l'application et une clé secrète.
- Vous avez copié le domaine AD de l'organisation.
- Les autorisations d'application suivantes ont été ajoutées à votre application AD sur l'option Microsoft Graph :
  - Lire des fichiers dans toutes les collections de sites (File.Read.All)
  - Lire le profil complet de tous les utilisateurs (User.Read.All)
  - Lire les données du répertoire (Directory.Read.All)
  - Lire tous les groupes (Group.Read.All)
  - Lire les éléments de toutes les collections du site (Site.Read.All)
- Copie de la liste des utilisateurs dont les documents doivent être indexés. Vous pouvez choisir de fournir une liste de noms d'utilisateur ou de fournir les noms d'utilisateur dans un fichier stocké dans un Amazon S3. Après avoir créé la source de données, vous pouvez :
  - Modifiez la liste des utilisateurs.
  - Passez d'une liste d'utilisateurs à une liste stockée dans un Amazon S3 bucket.
  - Modifiez l'emplacement du Amazon S3 compartiment d'une liste d'utilisateurs. Si vous modifiez l'emplacement du compartiment, vous devez également mettre à jour le IAM rôle de la source de données afin qu'elle ait accès au compartiment.

    > **Note**
    >
    > Si vous stockez la liste des noms d'utilisateur dans un Amazon S3 bucket, la IAM politique de la source de données doit fournir l'accès au bucket et l'accès à la clé avec laquelle le bucket a été chiffré, le cas échéant.
- Il est vérifié que chaque document est unique dans OneDrive et entre les autres sources de données que vous prévoyez d'utiliser pour le même index. Chaque source de données que vous souhaitez utiliser pour un index ne doit pas contenir le même document dans toutes les sources de données. IDs Les documents sont globaux par rapport à un index et doivent être uniques par index.

Dans votre Compte AWS, assurez-vous d'avoir :

- Création d'un Amazon Kendra index et, si vous utilisez l'API, notez l'ID de l'index.

- Vous avez créé un IAM rôle pour votre source de données et, si vous utilisez l'API, notez l'ARN du IAM rôle.

> **Note**
>
> Si vous modifiez votre type d'authentification et vos informations d'identification, vous devez mettre à jour votre IAM rôle pour accéder au bon identifiant AWS Secrets Manager secret.

- Stockez vos informations d' OneDrive authentification dans un AWS Secrets Manager secret et, si vous utilisez l'API, notez l'ARN du secret.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

Si vous n'avez pas de IAM rôle ou de secret existant, vous pouvez utiliser la console pour créer un nouveau IAM rôle et un nouveau Secrets Manager secret lorsque vous connectez votre source de OneDrive données à Amazon Kendra. Si vous utilisez l'API, vous devez fournir l'ARN d'un IAM rôle et d'un Secrets Manager secret existants, ainsi qu'un identifiant d'index.

Instructions de connexion

Pour vous connecter Amazon Kendra à votre source de OneDrive données, vous devez fournir les détails de vos OneDrive informations d'identification afin de Amazon Kendra pouvoir accéder à vos données. Si vous n'avez pas encore configuré OneDrive pour Amazon Kendra voirPrérequis.

Console

Pour vous connecter Amazon Kendra à OneDrive

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la.
2. Dans le volet de navigation de gauche, choisissez Index, puis choisissez l'index que vous souhaitez utiliser dans la liste des index.

> **Note**
>
> Vous pouvez choisir de configurer ou de modifier vos paramètres de contrôle d'accès utilisateur dans les paramètres de l'index.

3. Sur la page de démarrage, choisissez Ajouter une source de données.
4. Sur la page Ajouter une source de données, choisissez OneDrive connecteur, puis sélectionnez Ajouter un connecteur. Si vous utilisez la version 2 (le cas échéant), choisissez le OneDrive connecteur avec la balise « V2.0 ».
5. Sur la page Spécifier les détails de la source de données, entrez les informations suivantes :
   a. Dans Nom et description, pour Nom de la source de données : entrez le nom de votre source de données. Vous pouvez inclure des tirets, mais pas des espaces.
   b. (Facultatif) Description : entrez une description facultative pour votre source de données.
   c. Dans la langue par défaut : choisissez une langue pour filtrer vos documents pour l'index. Sauf indication contraire, la langue par défaut est l'anglais. La langue spécifiée dans les métadonnées du document remplace la langue sélectionnée.
   d. Dans Balises, pour Ajouter une nouvelle balise : incluez des balises facultatives pour rechercher et filtrer vos ressources ou suivre vos AWS coûts.
   e. Choisissez Suivant.
6. Sur la page Définir l'accès et la sécurité, entrez les informations suivantes :
   a. OneDrive ID du locataire —Entrez l'ID du OneDrive locataire sans le protocole.
   b. Type d'authentification : choisissez entre nouvelle et existante.
   c. i. Si vous choisissez Existant, sélectionnez un secret existant pour Sélectionner un secret.
      ii. Si vous choisissez Nouveau, entrez les informations suivantes dans la section Nouveau AWS Secrets Manager secret :
         A. Nom secret : nom de votre secret. Le préfixe « AmazonKendra - OneDrive - » est automatiquement ajouté à votre nom secret.
         B. Pour l'ID de l'application et le mot de passe de l'application : entrez les valeurs des informations d'authentification de votre OneDrive compte, puis choisissez Enregistrer l'authentification.

d. IAM rôle —Choisissez un IAM rôle existant ou créez-en un nouveau IAM pour accéder aux informations d'identification de votre référentiel et indexer le contenu.

> **Note**
>
> IAM les rôles utilisés pour les index ne peuvent pas être utilisés pour les sources de données. Si vous ne savez pas si un rôle existant est utilisé pour un index ou une FAQ, choisissez Créer un nouveau rôle pour éviter les erreurs.

e. Choisissez Suivant.

7. Sur la page Configurer les paramètres de synchronisation, entrez les informations suivantes :

a. Choisissez entre le fichier de liste et la liste de noms en fonction de votre cas d'utilisation.

i. Si vous choisissez Fichier de liste, entrez les informations suivantes :

- Sélectionnez l'emplacement —Entrez le chemin d'accès à votre Amazon S3 compartiment.

  Ajouter un fichier de liste d'utilisateurs à Amazon S3 —Sélectionnez cette option pour ajouter vos fichiers de liste d'utilisateurs à votre Amazon S3 compartiment.

  Mappages de groupes locaux d'utilisateurs : sélectionnez cette option pour utiliser le mappage de groupes locaux pour filtrer votre contenu.

ii. Si vous choisissez la liste des noms, entrez les informations suivantes :

- Nom d'utilisateur : entrez jusqu'à 10 lecteurs utilisateur à indexer. Pour ajouter plus de 10 utilisateurs, créez un fichier contenant les noms.

  Ajouter un autre : choisissez d'ajouter d'autres utilisateurs.

  Mappages de groupes locaux d'utilisateurs : sélectionnez cette option pour utiliser le mappage de groupes locaux pour filtrer votre contenu.

b. Pour les configurations supplémentaires : ajoutez des modèles d'expressions régulières pour inclure ou exclure certains fichiers. Vous pouvez ajouter jusqu'à 100 motifs.

c. Dans Synchroniser le calendrier d'exécution, pour Fréquence : choisissez la fréquence de synchronisation avec votre source de données. Amazon Kendra

d. Choisissez Suivant.

8. Sur la page Définir les mappages de champs, entrez les informations suivantes :

   a. Pour les champs de source de données par défaut et les mappages de champs suggérés supplémentaires, sélectionnez parmi les champs de source de données par défaut Amazon Kendra générés que vous souhaitez mapper à votre index.

   b. Choisissez Suivant.

9. Sur la page Réviser et créer, vérifiez que les informations saisies sont correctes, puis sélectionnez Ajouter une source de données. Vous pouvez également choisir de modifier vos informations à partir de cette page. Votre source de données apparaîtra sur la page Sources de données une fois que la source de données aura été ajoutée avec succès.

API

Pour vous connecter Amazon Kendra à OneDrive

Vous devez spécifier les éléments suivants à l'aide de l'OneDriveConfigurationAPI :

- ID du locataire : spécifiez le domaine Azure Active Directory de l'organisation.
- OneDrive Utilisateurs : spécifiez la liste des comptes utilisateurs dont les documents doivent être indexés.
- Nom de ressource Amazon (ARN) secret : indiquez le nom de ressource Amazon (ARN) d'un Secrets Manager secret contenant les informations d'authentification de votre OneDrive compte. Le secret est stocké dans une structure JSON avec les clés suivantes :

```
{
    "username": "OAuth client ID",
    "password": "client secret"
}
```

- IAM role —Spécifiez le `RoleArn` moment où vous appelez `CreateDataSource` pour fournir à un IAM rôle les autorisations d'accéder à votre Secrets Manager secret et d'appeler le public requis APIs pour le OneDrive connecteur et Amazon Kendra. Pour plus d'informations, consultez la section IAM Rôles des sources de OneDrive données.

Vous pouvez également ajouter les fonctionnalités optionnelles suivantes :

- Filtres d'inclusion et d'exclusion : indiquez si vous souhaitez inclure ou exclure certains documents.

  > **Note**
  >
  > La plupart des sources de données utilisent des modèles d'expressions régulières, qui sont des modèles d'inclusion ou d'exclusion appelés filtres. Si vous spécifiez un filtre d'inclusion, seul le contenu correspondant au filtre d'inclusion est indexé. Tout document qui ne correspond pas au filtre d'inclusion n'est pas indexé. Si vous spécifiez un filtre d'inclusion et d'exclusion, les documents correspondant au filtre d'exclusion ne sont pas indexés, même s'ils correspondent au filtre d'inclusion.

- Mappages de champs : choisissez de mapper les champs de votre source de OneDrive données à vos champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.

  > **Note**
  >
  > Le champ du corps du document ou l'équivalent du corps du document pour vos documents est requis pour Amazon Kendra effectuer une recherche dans vos documents. Vous devez associer le nom du champ du corps du document dans votre source de données au nom du champ d'index`_document_body`. Tous les autres champs sont facultatifs.

- Filtrage du contexte utilisateur et contrôle d'accèsAmazon Kendra : analyse la liste de contrôle d'accès (ACL) de vos documents, si vous disposez d'une ACL pour vos documents. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.

OneDrive Connecteur Microsoft V2.0

Microsoft OneDrive est un service de stockage basé sur le cloud que vous pouvez utiliser pour stocker, partager et héberger votre contenu. Vous pouvez l'utiliser Amazon Kendra pour indexer votre source de OneDrive données.

Vous pouvez vous connecter Amazon Kendra à votre source de OneDrive données à l'aide de la Amazon Kendra console et de l'OneDriveConfigurationAPI.

> **Note**
>
> Support pour OneDrive Connector V1.0/ OneDriveConfiguration API devrait prendre fin en juin 2023. Nous vous recommandons d'utiliser OneDrive Connector TemplateConfiguration V2.0/API. La version 2.0 fournit des fonctionnalités supplémentaires ACLs et des fonctionnalités d'explorateur d'identité.

Pour résoudre les problèmes liés à votre connecteur de source de Amazon Kendra OneDrive données, consultezDépannage des sources de données.

Rubriques

- Fonctionnalités prises en charge
- Prérequis
- Instructions de connexion

Fonctionnalités prises en charge

Amazon Kendra OneDrive le connecteur de source de données prend en charge les fonctionnalités suivantes :

- Mappages de champs
- Contrôle d'accès utilisateur
- Filtres d'inclusion/exclusion
- Synchronisation complète et incrémentielle du contenu
- Cloud privé virtuel (VPC)

Prérequis

Avant de pouvoir Amazon Kendra indexer votre source de OneDrive données, apportez ces modifications à vos AWS comptes OneDrive and.

Dans OneDrive, assurez-vous d'avoir :

- Vous avez créé un OneDrive compte dans Office 365.

- Vous avez noté votre identifiant de client Microsoft 365. Vous pouvez trouver votre ID de locataire dans les propriétés de votre portail Azure Active Directory ou dans votre OAuth application.
- Vous avez créé une OAuth application sur le portail Azure et avez noté l'ID client et le secret du client ou les informations d'identification du client utilisées pour l'authentification avec un AWS Secrets Manager secret. Consultez le didacticiel Microsoft et l'exemple d'application enregistrée pour plus d'informations.

> **Note**
>
> Lorsque vous créez ou enregistrez une application sur le portail Azure, l'ID secret représente la valeur secrète réelle. Vous devez prendre note ou enregistrer la valeur secrète réelle immédiatement lors de la création du secret et de l'application. Vous pouvez accéder à votre secret en sélectionnant le nom de votre application sur le portail Azure, puis en accédant à l'option de menu sur les certificats et les secrets.
> Vous pouvez accéder à votre ID client en sélectionnant le nom de votre application sur le portail Azure, puis en accédant à la page de présentation. L'ID de l'application (client) est l'ID du client.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

- A utilisé l'ID de l'application AD pour enregistrer une clé secrète pour l'application sur le site AD. La clé secrète doit contenir l'ID de l'application et une clé secrète.
- Vous avez copié le domaine AD de l'organisation.
- Vous avez ajouté les autorisations suivantes à votre application AD sur l'option Microsoft Graph :
  - Lire des fichiers dans toutes les collections de sites (File.Read.All)
  - Lire les profils complets de tous les utilisateurs (User.Read.All)
  - Lire tous les groupes (Group.Read.All)
  - Lire toutes les notes (Notes.Read.All)

- Copie de la liste des utilisateurs dont les documents doivent être indexés. Vous pouvez choisir de fournir une liste de noms d'utilisateur ou de fournir les noms d'utilisateur dans un fichier stocké dans un Amazon S3. Après avoir créé la source de données, vous pouvez :
  - Modifiez la liste des utilisateurs.
  - Passez d'une liste d'utilisateurs à une liste stockée dans un Amazon S3 bucket.
  - Modifiez l'emplacement du Amazon S3 compartiment d'une liste d'utilisateurs. Si vous modifiez l'emplacement du compartiment, vous devez également mettre à jour le IAM rôle de la source de données afin qu'elle ait accès au compartiment.

    > **Note**
    >
    > Si vous stockez la liste des noms d'utilisateur dans un Amazon S3 bucket, la IAM politique de la source de données doit fournir l'accès au bucket et l'accès à la clé avec laquelle le bucket a été chiffré, le cas échéant.
    > Le OneDrive connecteur utilise le courrier électronique provenant des informations de contact présentes dans les propriétés utilisateur de Onedrive. Assurez-vous que le champ e-mail de l'utilisateur dont vous souhaitez analyser les données est configuré dans la page Informations de contact, car pour les nouveaux utilisateurs, ce champ peut être vide.

Dans votre AWS compte, assurez-vous d'avoir :

- Création d'un Amazon Kendra index et, si vous utilisez l'API, notez l'identifiant de l'index.
- Vous avez créé un IAM rôle pour votre source de données et, si vous utilisez l'API, notez l'ARN du IAM rôle.
- Stockez vos informations d' OneDrive authentification dans un AWS Secrets Manager secret et, si vous utilisez l'API, notez l'ARN du secret.

Si vous n'avez pas de IAM rôle ou de secret existant, vous pouvez utiliser la console pour créer un nouveau IAM rôle et un nouveau Secrets Manager secret lorsque vous connectez votre source de OneDrive données à Amazon Kendra. Si vous utilisez l'API, vous devez fournir l'ARN d'un IAM rôle et d'un Secrets Manager secret existants, ainsi qu'un identifiant d'index.

Instructions de connexion

Pour vous connecter Amazon Kendra à votre source de OneDrive données, vous devez fournir les détails de vos OneDrive informations d'identification afin de Amazon Kendra pouvoir accéder à vos données. Si vous n'avez pas encore configuré OneDrive pour Amazon Kendra, consultezPrérequis.

Console

Pour vous connecter Amazon Kendra à OneDrive

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la.
2. Dans le volet de navigation de gauche, choisissez Index, puis choisissez l'index que vous souhaitez utiliser dans la liste des index.

> **Note**
>
> Vous pouvez choisir de configurer ou de modifier vos paramètres de contrôle d'accès utilisateur dans les paramètres de l'index.

3. Sur la page de démarrage, choisissez Ajouter une source de données.
4. Sur la page Ajouter une source de données, choisissez OneDrive connecteur, puis sélectionnez Ajouter un connecteur. Si vous utilisez la version 2 (le cas échéant), choisissez le OneDrive connecteur avec la balise « V2.0 ».
5. Sur la page Spécifier les détails de la source de données, entrez les informations suivantes :
   a. Dans Nom et description, pour Nom de la source de données : entrez le nom de votre source de données. Vous pouvez inclure des tirets, mais pas des espaces.
   b. (Facultatif) Description : entrez une description facultative pour votre source de données.
   c. Dans la langue par défaut : choisissez une langue pour filtrer vos documents pour l'index. Sauf indication contraire, la langue par défaut est l'anglais. La langue spécifiée dans les métadonnées du document remplace la langue sélectionnée.
   d. Dans Balises, pour Ajouter une nouvelle balise : incluez des balises facultatives pour rechercher et filtrer vos ressources ou suivre vos AWS coûts.
   e. Choisissez Suivant.
6. Sur la page Définir l'accès et la sécurité, entrez les informations suivantes :
   a. OneDrive ID du locataire —Entrez l'ID du OneDrive locataire sans le protocole.

b. Autorisation : activez ou désactivez les informations de la liste de contrôle d'accès (ACL) pour vos documents, si vous disposez d'une ACL et que vous souhaitez l'utiliser pour le contrôle d'accès. L'ACL indique les documents auxquels les utilisateurs et les groupes peuvent accéder. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.

c. Dans Authentification, choisissez entre nouveau et existant.

d. i. Si vous choisissez Existant, sélectionnez un secret existant pour Sélectionner un secret.

   ii. Si vous choisissez Nouveau, entrez les informations suivantes dans la section Nouveau AWS Secrets Manager secret :

      A. Nom secret : nom de votre secret. Le préfixe « AmazonKendra - OneDrive - » est automatiquement ajouté à votre nom secret.

      B. Pour l'ID client et le secret du client : entrez l'ID client et le secret du client.

e. Virtual Private Cloud (VPC) —Vous pouvez choisir d'utiliser un VPC. Dans ce cas, vous devez ajouter des sous-réseaux et des groupes de sécurité VPC.

f. Identity Crawler : spécifiez s'il faut activer l'explorateur Amazon Kendra d'identité. Le robot d'exploration d'identité utilise les informations de la liste de contrôle d'accès (ACL) de vos documents pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Si vous disposez d'une ACL pour vos documents et que vous choisissez de l'utiliser, vous pouvez également choisir d'activer le robot d'exploration Amazon Kendra d'identité pour configurer le filtrage des résultats de recherche par contexte utilisateur. Dans le cas contraire, si Identity Crawler est désactivé, tous les documents peuvent faire l'objet d'une recherche publique. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'exploration d'identité est désactivé, vous pouvez également utiliser l'PutPrincipalMappingAPI pour télécharger les informations d'accès des utilisateurs et des groupes afin de filtrer le contexte utilisateur.

g. IAM rôle —Choisissez un IAM rôle existant ou créez-en un nouveau IAM pour accéder aux informations d'identification de votre référentiel et indexer le contenu.

> **Note**
>
> IAM les rôles utilisés pour les index ne peuvent pas être utilisés pour les sources de données. Si vous ne savez pas si un rôle existant est utilisé pour un index ou une FAQ, choisissez Créer un nouveau rôle pour éviter les erreurs.

h. Choisissez Suivant.

7. Sur la page Configurer les paramètres de synchronisation, entrez les informations suivantes :

8. a. Pour l'étendue de la synchronisation : choisissez les OneDrive données des utilisateurs à indexer. Vous pouvez ajouter un maximum de 10 utilisateurs manuellement.

   b. Pour les configurations supplémentaires : ajoutez des modèles d'expressions régulières pour inclure ou exclure certains contenus. Vous pouvez ajouter jusqu'à 100 motifs.

   c. Mode de synchronisation : choisissez la manière dont vous souhaitez mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation.

      - Synchronisation complète : Indexez tout le contenu fraîchement, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
      - Nouvelle synchronisation modifiée : Indexez uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
      - Synchronisation nouvelle, modifiée ou supprimée : Indexez uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

   d. Dans Calendrier d'exécution de la synchronisation, pour Fréquence : choisissez la fréquence à laquelle vous souhaitez synchroniser le contenu de votre source de données et mettre à jour votre index.

e. Choisissez Suivant.

9. Sur la page Définir les mappages de champs, entrez les informations suivantes :

   a. Champs de source de données par défaut : sélectionnez parmi les champs de source de données par défaut Amazon Kendra générés que vous souhaitez mapper à votre index.

   b. Choisissez Suivant.

10. Sur la page Réviser et créer, vérifiez que les informations saisies sont correctes, puis sélectionnez Ajouter une source de données. Vous pouvez également choisir de modifier vos informations à partir de cette page. Votre source de données apparaîtra sur la page Sources de données une fois que la source de données aura été ajoutée avec succès.

API

Pour vous connecter Amazon Kendra à OneDrive

Vous devez spécifier un JSON du schéma de source de données à l'aide de l'TemplateConfigurationAPI. Vous devez fournir les informations suivantes :

- Source de données —Spécifiez le type de source de données tel que `ONEDRIVEV2` lorsque vous utilisez le schéma TemplateConfigurationJSON. Spécifiez également la source de données `TEMPLATE` lorsque vous appelez l'CreateDataSourceAPI.
- ID de locataire —Spécifiez l'ID de locataire Microsoft 365. Vous pouvez trouver votre ID de locataire dans les propriétés de votre portail Azure Active Directory ou dans votre OAuth application.
- Mode de synchronisation : spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation. Vous pouvez choisir entre :
  - `FORCED_FULL_CRAWL`pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
  - `FULL_CRAWL`pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

  - `CHANGE_LOG`pour indexer uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
- Nom de ressource Amazon (ARN) secret : indiquez le nom de ressource Amazon (ARN) d'un Secrets Manager secret contenant les informations d'authentification que vous avez créées dans votre OneDrive compte.

  Si vous utilisez l'authentification OAuth 2.0, le secret est stocké dans une structure JSON avec les clés suivantes :

```
{
    "clientId": "client ID",
    "clientSecret": "client secret"
}
```

- IAM role —Spécifiez le `RoleArn` moment où vous appelez `CreateDataSource` pour fournir à un IAM rôle les autorisations d'accéder à votre Secrets Manager secret et d'appeler le public requis APIs pour le OneDrive connecteur et Amazon Kendra. Pour plus d'informations, consultez la section IAM Rôles des sources de OneDrive données.

Vous pouvez également ajouter les fonctionnalités optionnelles suivantes :

- Virtual Private Cloud (VPC) **`VpcConfiguration`** —Spécifiez le moment de votre appel. `CreateDataSource` Pour de plus amples informations, veuillez consulter Configuration Amazon Kendra pour utiliser un Amazon VPC.
- Filtres d'inclusion et d'exclusion : vous pouvez indiquer si vous souhaitez inclure ou exclure certains fichiers, OneNote sections et OneNote pages.

> **Note**
>
> La plupart des sources de données utilisent des modèles d'expressions régulières, qui sont des modèles d'inclusion ou d'exclusion appelés filtres. Si vous spécifiez un filtre d'inclusion, seul le contenu correspondant au filtre d'inclusion est indexé. Tout document qui ne correspond pas au filtre d'inclusion n'est pas indexé. Si vous spécifiez un filtre d'inclusion et d'exclusion, les documents correspondant au filtre d'exclusion ne sont pas indexés, même s'ils correspondent au filtre d'inclusion.

- Identity Crawler : spécifiez s'il faut activer l'explorateur Amazon Kendra d'identité. Le robot d'exploration d'identité utilise les informations de la liste de contrôle d'accès (ACL) de vos documents pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Si vous disposez d'une ACL pour vos documents et que vous choisissez de l'utiliser, vous pouvez également choisir d'activer le robot d'exploration Amazon Kendra d'identité pour configurer le filtrage des résultats de recherche par contexte utilisateur. Dans le cas contraire, si Identity Crawler est désactivé, tous les documents peuvent faire l'objet d'une recherche publique. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'exploration d'identité est désactivé, vous pouvez également utiliser l'PutPrincipalMappingAPI pour télécharger les informations d'accès des utilisateurs et des groupes afin de filtrer le contexte utilisateur.
- Mappages de champs : vous ne pouvez mapper que des champs d'index intégrés ou communs pour le Amazon Kendra OneDrive connecteur. Le mappage de champs personnalisé n'est pas disponible pour le OneDrive connecteur en raison des limites de l'API. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.

Pour obtenir la liste des autres clés JSON importantes à configurer, consultez le schéma du OneDrive modèle.

En savoir plus

Pour en savoir plus sur l'intégration Amazon Kendra à votre source de OneDrive données, consultez :

- Annonce de la mise à jour OneDrive du connecteur Microsoft (V2) pour Amazon Kendra.

Remarques

- Lorsque les listes de contrôle d'accès (ACLs) sont activées, l'option « Synchroniser uniquement le contenu nouveau ou modifié » n'est pas disponible en raison des limites de OneDrive l'API. Nous vous recommandons d'utiliser plutôt les modes « Synchronisation complète » ou « Synchronisation du contenu nouveau, modifié ou supprimé », ou de les désactiver ACLs si vous devez utiliser ce mode de synchronisation.

Microsoft SharePoint

SharePoint est un service collaboratif de création de sites Web que vous pouvez utiliser pour personnaliser le contenu Web et créer des pages, des sites, des bibliothèques de documents et des listes. Vous pouvez l'utiliser Amazon Kendra pour indexer votre source de SharePoint données.

Amazon Kendra est actuellement compatible avec SharePoint Online et SharePoint Server (versions 2013, 2016, 2019 et édition par abonnement).

Vous pouvez vous connecter Amazon Kendra à votre source de SharePoint données à l'aide de la Amazon Kendra console, de l'TemplateConfigurationAPI ou de l'SharePointConfigurationAPI.

Amazon Kendra possède deux versions du SharePoint connecteur. Les fonctionnalités prises en charge par chaque version incluent :

SharePoint Connecteur V1.0/API SharePointConfiguration

- Mappages de champs
- Contrôle d'accès utilisateur
- Filtres d'inclusion/exclusion
- Journal des modifications
- Cloud privé virtuel (VPC)

SharePoint Connecteur V2.0/API TemplateConfiguration

- Mappages de champs
- Contrôle d'accès utilisateur
- Filtres d'inclusion/exclusion
- Synchronisation complète et incrémentielle du contenu
- Cloud privé virtuel (VPC)

> **Note**
>
> SharePoint le connecteur V1.0/ SharePointConfiguration API a pris fin en 2023. Nous vous recommandons de migrer vers ou d'utiliser le SharePoint connecteur TemplateConfiguration V2.0/API.

Pour résoudre les problèmes liés à votre connecteur de source de Amazon Kendra SharePoint données, consultezDépannage des sources de données.

Rubriques

- SharePoint connecteur V1.0
- SharePoint connecteur V2.0

SharePoint connecteur V1.0

SharePoint est un service collaboratif de création de sites Web que vous pouvez utiliser pour personnaliser le contenu Web et créer des pages, des sites, des bibliothèques de documents et des listes. Si vous êtes un SharePoint utilisateur, vous pouvez l'utiliser Amazon Kendra pour indexer votre source de SharePoint données.

> **Note**
>
> SharePoint le connecteur V1.0/ SharePointConfiguration API a pris fin en 2023. Nous vous recommandons de migrer vers ou d'utiliser le SharePoint connecteur TemplateConfiguration V2.0/API.

Pour résoudre les problèmes liés à votre connecteur de source de Amazon Kendra SharePoint données, consultezDépannage des sources de données.

Rubriques

- Fonctionnalités prises en charge
- Prérequis
- Instructions de connexion
- En savoir plus

Fonctionnalités prises en charge

- Mappages de champs
- Contrôle d'accès utilisateur
- Filtres d'inclusion/exclusion

- Journal des modifications
- Cloud privé virtuel (VPC)

Prérequis

Avant de pouvoir Amazon Kendra indexer votre source de SharePoint données, apportez ces modifications à vos AWS comptes SharePoint and.

Vous devez fournir des informations d'authentification, que vous stockez en toute sécurité dans un AWS Secrets Manager secret.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

Dans SharePoint, assurez-vous d'avoir :

- Notez l'URL des SharePoint sites que vous souhaitez indexer.
- Pour SharePoint en ligne :
  - Vous avez noté vos informations d'authentification de base contenant un nom d'utilisateur et un mot de passe avec des autorisations d'administrateur du site.
  - Facultatif : informations d'identification OAuth 2.0 générées contenant un nom d'utilisateur, un mot de passe, un identifiant client et un secret client.
  - Paramètres de sécurité par défaut désactivés sur votre portail Azure à l'aide d'un utilisateur administratif. Pour plus d'informations sur la gestion des paramètres de sécurité par défaut sur le portail Azure, consultez la documentation Microsoft sur la procédure à suivre pour définir les paramètres enable/disable de sécurité par défaut.
- Pour le SharePoint serveur :
  - Vous avez noté le nom de domaine de votre SharePoint serveur (le nom NetBIOS dans votre Active Directory). Vous l'utilisez, ainsi que votre nom d'utilisateur et votre mot de passe d'authentification de SharePoint base, pour connecter le SharePoint serveur à Amazon Kendra.

> **Note**
>
> Si vous utilisez SharePoint Server et devez convertir votre liste de contrôle d'accès (ACL) au format e-mail pour le filtrage en fonction du contexte utilisateur, fournissez l'URL du serveur LDAP et la base de recherche LDAP. Vous pouvez également utiliser le remplacement du domaine de l'annuaire. L'URL du serveur LDAP est le nom de domaine complet et le numéro de port (par exemple, ldap : //example.com:389). La base de recherche LDAP est constituée des contrôleurs de domaine « example » et « com ». Avec le remplacement du domaine de l'annuaire, vous pouvez utiliser le domaine de messagerie au lieu d'utiliser l'URL du serveur LDAP et la base de recherche LDAP. Par exemple, le domaine de messagerie pour username@example.com est « exemple.com ». Vous pouvez utiliser cette dérogation si la validation de votre domaine ne vous préoccupe pas et si vous souhaitez simplement utiliser votre domaine de messagerie.

- Vous avez ajouté les autorisations suivantes à votre SharePoint compte :

  Pour les SharePoint listes

  - Éléments ouverts : affichez la source des documents à l'aide des gestionnaires de fichiers côté serveur.
  - Afficher les pages de candidature : affichez les formulaires, les vues et les pages de candidature. Énumérez des listes.
  - Afficher les éléments : affichez les éléments dans les listes et les documents dans les bibliothèques de documents.
  - Afficher les versions : permet d'afficher les versions précédentes d'un élément de liste ou d'un document.

  Pour les SharePoint sites Web

  - Parcourir les répertoires : énumère les fichiers et les dossiers d'un site Web à l'aide de SharePoint Designer et de l'interface Web DAV.
  - Parcourir les informations utilisateur : affichez les informations sur les utilisateurs du site Web.
  - Énumérer les autorisations : énumère les autorisations sur le site Web, la liste, le dossier, le document ou l'élément de liste.
  - Ouvrir : ouvrez un site Web, une liste ou un dossier pour accéder aux éléments contenus dans le conteneur.

  - Utiliser les fonctionnalités d'intégration du client : utilisez SOAP, WebDAV, le modèle d'objet client ou SharePoint les interfaces Designer pour accéder au site Web.
  - Utiliser des interfaces distantes : utilisez des fonctionnalités qui lancent des applications clientes.
  - Afficher les pages : affiche les pages d'un site Web.
- Il est vérifié que chaque document est unique dans SharePoint et entre les autres sources de données que vous prévoyez d'utiliser pour le même index. Chaque source de données que vous souhaitez utiliser pour un index ne doit pas contenir le même document dans toutes les sources de données. IDs Les documents sont globaux par rapport à un index et doivent être uniques par index.

Dans votre Compte AWS, assurez-vous d'avoir :

- Création d'un Amazon Kendra index et, si vous utilisez l'API, notez l'ID de l'index.
- Vous avez créé un IAM rôle pour votre source de données et, si vous utilisez l'API, notez l'ARN du IAM rôle.

> **Note**
>
> Si vous modifiez votre type d'authentification et vos informations d'identification, vous devez mettre à jour votre IAM rôle pour accéder au bon identifiant AWS Secrets Manager secret.

- Stockez vos informations d' SharePoint authentification dans un AWS Secrets Manager secret et, si vous utilisez l'API, notez l'ARN du secret.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

Si vous n'avez pas de IAM rôle ou de secret existant, vous pouvez utiliser la console pour créer un nouveau IAM rôle et un nouveau Secrets Manager secret lorsque vous connectez votre source de

SharePoint données à Amazon Kendra. Si vous utilisez l'API, vous devez fournir l'ARN d'un IAM rôle et d'un Secrets Manager secret existants, ainsi qu'un identifiant d'index.

Instructions de connexion

Pour vous connecter Amazon Kendra à votre source de SharePoint données, vous devez fournir les détails de vos SharePoint informations d'identification afin de Amazon Kendra pouvoir accéder à vos données. Si vous n'avez pas encore configuré SharePoint pour Amazon Kendra voirPrérequis.

Console

Pour vous connecter Amazon Kendra à SharePoint

1. Connectez-vous à la console AWS de gestion et Amazon Kendra ouvrez-la.
2. Dans le volet de navigation de gauche, choisissez Index, puis choisissez l'index que vous souhaitez utiliser dans la liste des index.

   > **Note**
   >
   > Vous pouvez choisir de configurer ou de modifier vos paramètres de contrôle d'accès utilisateur dans les paramètres de l'index.

3. Sur la page de démarrage, choisissez Ajouter une source de données.
4. Sur la page Ajouter une source de données, choisissez le SharePoint connecteur v1.0, puis sélectionnez Ajouter une source de données.
5. Sur la page Spécifier les détails de la source de données, entrez les informations suivantes :

   a. Dans Nom et description, pour Nom de la source de données : entrez le nom de votre source de données. Vous pouvez inclure des tirets, mais pas des espaces.

   b. (Facultatif) Description : entrez une description facultative pour votre source de données.

   c. Dans la langue par défaut : choisissez une langue pour filtrer vos documents pour l'index. Sauf indication contraire, la langue par défaut est l'anglais. La langue spécifiée dans les métadonnées du document remplace la langue sélectionnée.

   d. Dans Balises, pour Ajouter une nouvelle balise : incluez des balises facultatives pour rechercher et filtrer vos ressources ou suivre vos AWS coûts.

   e. Choisissez Suivant.

6. Sur la page Définir l'accès et la sécurité, entrez les informations suivantes :

a. Pour la méthode d'hébergement : choisissez entre SharePoint Online et SharePointServer.

   i. Pour SharePointen ligne —Entrez le site URLs spécifique à votre SharePoint référentiel.

   ii. Pour le SharePointserveur : choisissez votre SharePoint version, entrez le site URLs spécifique à votre SharePoint référentiel, puis le Amazon S3 chemin d'accès à l'emplacement de votre certificat SSL.

b. (SharePoint Serveur uniquement) Pour le proxy Web : entrez le nom d'hôte et le numéro de port de votre SharePoint instance interne. Le numéro de port doit être une valeur numérique comprise entre 0 et 65535.

c. Pour l'authentification : choisissez l'une des options suivantes en fonction de votre cas d'utilisation :

   i. Pour SharePoint en ligne : choisissez entre l'authentification de base et l'authentification OAuth 2.0.

   ii. Pour le SharePoint serveur : choisissez entre None, LDAP et Manual.

d. Pour le AWS Secrets Manager secret : choisissez un secret existant ou créez-en un nouveau Secrets Manager pour stocker vos informations SharePoint d'authentification. Si vous choisissez de créer un nouveau secret, une fenêtre AWS Secrets Manager secrète s'ouvre. Vous devez saisir un nom secret. Le préfixe « AmazonKendra - SharePoint - » est automatiquement ajouté à votre nom secret.

e. Entrez les autres informations suivantes dans la fenêtre Créer un AWS Secrets Manager secret :

   i. Choisissez parmi les options d'authentification SharePoint dans le cloud suivantes, en fonction de votre cas d'utilisation :

      A. Authentification de base —Entrez le nom d'utilisateur de votre SharePoint compte comme nom d'utilisateur et le mot de passe du SharePoint compte comme mot de passe.

      B. OAuth Authentification 2.0 —Entrez le nom d'utilisateur de votre SharePoint compte comme nom d'utilisateur, le mot de passe du SharePoint compte comme mot de passe, votre SharePoint identifiant unique généré automatiquement comme identifiant client et la chaîne secrète partagée utilisée par les deux SharePoint et Amazon Kendra comme secret client.

ii. Choisissez l'une des options d'authentification SharePoint du serveur suivantes, en fonction de votre cas d'utilisation :

A. Aucun —Entrez le nom d'utilisateur de votre SharePoint compte comme nom d'utilisateur, le mot de passe de votre SharePoint compte comme mot de passe et le nom de domaine de votre serveur.

B. LDAP —Entrez le nom d'utilisateur de votre SharePoint compte comme nom d'utilisateur, le mot de passe du SharePoint compte comme mot de passe, le point de terminaison de votre serveur LDAP (y compris le protocole et le numéro de port, par exemple`ldap://example.com:389`) et votre base de recherche LDAP (par exemple,). `dc=example, dc=com`

C. Manuel —Entrez le nom d'utilisateur de votre SharePoint compte en tant que nom d'utilisateur, le mot de passe de votre SharePoint compte en tant que mot de passe et votre remplacement de domaine de messagerie (domaine de messagerie de l'utilisateur ou du groupe de l'annuaire).

iii. Choisissez Enregistrer.

f. Virtual Private Cloud (VPC) : vous devez également ajouter des sous-réseaux et des groupes de sécurité VPC.

> **Note**
>
> Vous devez utiliser un VPC si vous utilisez SharePoint un serveur. Amazon VPC est facultatif pour les autres SharePoint versions.

g. IAM rôle —Choisissez un IAM rôle existant ou créez-en un nouveau IAM pour accéder aux informations d'identification de votre référentiel et indexer le contenu.

> **Note**
>
> IAM les rôles utilisés pour les index ne peuvent pas être utilisés pour les sources de données. Si vous ne savez pas si un rôle existant est utilisé pour un index ou une FAQ, choisissez Créer un nouveau rôle pour éviter les erreurs.

h. Choisissez Suivant.

7. Sur la page Configurer les paramètres de synchronisation, entrez les informations suivantes :

a. Utiliser le journal des modifications : sélectionnez cette option pour mettre à jour votre index au lieu de synchroniser tous vos fichiers.

b. Analyser les pièces jointes : sélectionnez cette option pour analyser les pièces jointes.

c. Utiliser des mappages de groupes locaux : sélectionnez cette option pour vous assurer que les documents sont correctement filtrés.

d. Configuration supplémentaire —Ajoutez des modèles d'expressions régulières pour inclure ou exclure certains fichiers. Vous pouvez ajouter jusqu'à 100 motifs.

e. Dans Synchroniser le calendrier d'exécution pour la fréquence : à quelle fréquence Amazon Kendra se synchronisera avec votre source de données.

f. Choisissez Suivant.

8. Sur la page Définir les mappages de champs, entrez les informations suivantes :

a. Amazon Kendra mappages de champs par défaut : sélectionnez parmi les champs de source de données par défaut Amazon Kendra générés que vous souhaitez mapper à votre index.

b. Pour les mappages de champs personnalisés : ajoutez des champs de source de données personnalisés pour créer un nom de champ d'index auquel mapper et le type de données du champ.

c. Choisissez Suivant.

9. Sur la page Réviser et créer, vérifiez que les informations saisies sont correctes, puis sélectionnez Ajouter une source de données. Vous pouvez également choisir de modifier vos informations à partir de cette page. Votre source de données apparaîtra sur la page Sources de données une fois que la source de données aura été ajoutée avec succès.

API

Pour vous connecter Amazon Kendra à SharePoint

Vous devez spécifier les éléments suivants à l'aide de l'SharePointConfigurationAPI :

- SharePointVersion : spécifiez la SharePoint version que vous utilisez lors de la configuration SharePoint. C'est le cas, que vous utilisiez SharePoint Server 2013, SharePoint Server 2016, SharePoint Server 2019 ou SharePoint Online.

- Nom de ressource Amazon (ARN) secret : indiquez le nom de ressource Amazon (ARN) d'un Secrets Manager secret contenant les informations d'authentification que vous avez créées dans votre SharePoint compte. Le secret est stocké dans une structure JSON.

  Pour l'authentification de base SharePoint en ligne, voici la structure JSON minimale qui doit figurer dans votre secret :

  ```
  {
      "userName": "user name",
      "password": "password"
  }
  ```

  Pour l'authentification SharePoint en ligne OAuth 2.0, voici la structure JSON minimale qui doit figurer dans votre secret :

  ```
  {
      "userName": "SharePoint account user name"",
      "password": "SharePoint account password",
      "clientId": "SharePoint auto-generated unique client id",
      "clientSecret": "secret string shared by Amazon Kendra and SharePoint to
   authorize communications"
  }
  ```

  Pour l'authentification de base du SharePoint serveur, voici la structure JSON minimale qui doit figurer dans votre secret :

  ```
  {
      "userName": "user name",
      "password": "password",
      "domain": "server domain name"
  }
  ```

  Pour l'authentification LDAP SharePoint du serveur (si vous devez convertir votre liste de contrôle d'accès (ACL) au format e-mail pour filtrer le contexte utilisateur, vous pouvez inclure l'URL du serveur LDAP et la base de recherche LDAP dans votre secret), voici la structure JSON minimale qui doit figurer dans votre secret :

  ```
  {
      "userName": "user name",
      "password": "password",
  ```

```
    "domain": "server domain name"
    "ldapServerUrl": "ldap://example.com:389",
    "ldapSearchBase": "dc=example,dc=com"
}
```

Pour l'authentification manuelle du SharePoint serveur, voici la structure JSON minimale qui doit figurer dans votre secret :

```
{
    "userName": "user name",
    "password": "password",
    "domain": "server domain name",
    "emailDomainOverride": "example.com"
}
```

- IAM role —Spécifiez le `RoleArn` moment où vous appelez `CreateDataSource` pour fournir à un IAM rôle les autorisations d'accéder à votre Secrets Manager secret et d'appeler le public requis APIs pour le SharePoint connecteur et Amazon Kendra. Pour plus d'informations, consultez la section IAM Rôles des sources de SharePoint données.
- Amazon VPC—Si vous utilisez SharePoint Server, spécifiez-le dans `VpcConfiguration` le cadre de la configuration de la source de données. Reportez-vous Amazon Kendra à la section Configuration pour utiliser un VPC.

Vous pouvez également ajouter les fonctionnalités optionnelles suivantes :

- Proxy Web : s'il faut se connecter à votre SharePoint site URLs via un proxy Web. Vous ne pouvez utiliser cette option que pour le SharePoint serveur.
- Listes d'indexation : indique si le contenu des pièces jointes Amazon Kendra doit être indexé sur les éléments de la SharePoint liste.
- Journal des modifications : Amazon Kendra faut-il utiliser le mécanisme du journal des modifications de la source de SharePoint données pour déterminer si un document doit être mis à jour dans l'index.

> **Note**
>
> Utilisez le journal des modifications si vous ne Amazon Kendra souhaitez pas numériser tous les documents. Si votre journal des modifications est volumineux, la numérisation des documents de la source de SharePoint données peut prendre Amazon Kendra

moins de temps que le traitement du journal des modifications. Si vous synchronisez votre source de SharePoint données avec votre index pour la première fois, tous les documents sont numérisés.

- Filtres d'inclusion et d'exclusion : vous pouvez indiquer si vous souhaitez inclure ou exclure certains contenus.

  > **Note**
  >
  > La plupart des sources de données utilisent des modèles d'expressions régulières, qui sont des modèles d'inclusion ou d'exclusion appelés filtres. Si vous spécifiez un filtre d'inclusion, seul le contenu correspondant au filtre d'inclusion est indexé. Tout document qui ne correspond pas au filtre d'inclusion n'est pas indexé. Si vous spécifiez un filtre d'inclusion et d'exclusion, les documents correspondant au filtre d'exclusion ne sont pas indexés, même s'ils correspondent au filtre d'inclusion.

- Mappages de champs : choisissez de mapper les champs de votre source de SharePoint données à vos champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.

  > **Note**
  >
  > Le champ du corps du document ou l'équivalent du corps du document pour vos documents est requis pour Amazon Kendra effectuer une recherche dans vos documents. Vous devez associer le nom du champ du corps du document dans votre source de données au nom du champ d'index`_document_body`. Tous les autres champs sont facultatifs.

- Filtrage du contexte utilisateur et contrôle d'accèsAmazon Kendra : analyse la liste de contrôle d'accès (ACL) de vos documents, si vous disposez d'une ACL pour vos documents. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.

En savoir plus

Pour en savoir plus sur l'intégration Amazon Kendra à votre source de SharePoint données, consultez :

- Commencer à utiliser le connecteur Amazon Kendra SharePoint en ligne

SharePoint connecteur V2.0

SharePoint est un service collaboratif de création de sites Web que vous pouvez utiliser pour personnaliser le contenu Web et créer des pages, des sites, des bibliothèques de documents et des listes. Vous pouvez l'utiliser Amazon Kendra pour indexer votre source de SharePoint données.

Amazon Kendra est actuellement compatible avec SharePoint Online et SharePoint Server (2013, 2016, 2019 et édition par abonnement).

> **Note**
>
> SharePoint le connecteur V1.0/ SharePointConfiguration API a pris fin en 2023. Nous vous recommandons de migrer vers ou d'utiliser le SharePoint connecteur TemplateConfiguration V2.0/API.

Pour résoudre les problèmes liés à votre connecteur de source de Amazon Kendra SharePoint données, consultezDépannage des sources de données.

Rubriques

- Fonctionnalités prises en charge
- Prérequis
- Instructions de connexion
- Remarques

Fonctionnalités prises en charge

Amazon Kendra SharePoint le connecteur de source de données prend en charge les fonctionnalités suivantes :

- Mappages de champs
- Contrôle d'accès utilisateur
- Filtres d'inclusion/exclusion
- Synchronisation complète et incrémentielle du contenu
- Cloud privé virtuel (VPC)

Prérequis

Avant de pouvoir Amazon Kendra indexer votre source de SharePoint données, apportez ces modifications à vos AWS comptes SharePoint and.

Vous devez fournir des informations d'authentification, que vous stockez en toute sécurité dans un AWS Secrets Manager secret.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

Dans SharePoint Online, assurez-vous d'avoir :

- Vous avez copié votre SharePoint instance URLs. Le format de l'URL d'hôte que vous entrez est*`https://yourdomain.com/sites/mysite`*. Votre URL doit commencer par`https`.
- Vous avez copié le nom de domaine de l'URL de votre SharePoint instance.
- Vous avez noté vos informations d'authentification de base contenant le nom d'utilisateur et le mot de passe avec les autorisations d'administrateur du site pour vous connecter à SharePoint Online.
- Paramètres de sécurité par défaut désactivés sur votre portail Azure à l'aide d'un utilisateur administratif. Pour plus d'informations sur la gestion des paramètres de sécurité par défaut sur le portail Azure, consultez la documentation Microsoft sur la procédure à suivre pour définir les paramètres enable/disable de sécurité par défaut.
- L'authentification multifactorielle (MFA) a été désactivée dans votre SharePoint compte, afin de ne pas empêcher l' Amazon Kendra exploration de votre contenu. SharePoint
- Si vous utilisez un type d'authentification autre que l'authentification de base : copiez l'ID de locataire de votre SharePoint instance. Pour plus de détails sur la façon de trouver votre identifiant de locataire, voir Rechercher votre identifiant de locataire Microsoft 365.
- Si vous devez passer à l'authentification des utilisateurs dans le cloud avec Microsoft Entra, consultez la documentation Microsoft sur l'authentification dans le cloud.
- Pour l'authentification OAuth OAuth 2.0 et l'authentification par jeton d'actualisation 2.0 : notez vos informations d'authentification de base contenant le nom d'utilisateur et le mot de passe que vous

utilisez pour vous connecter à SharePoint Online, ainsi que l'ID client et le secret du client générés après l'enregistrement SharePoint auprès d'Azure AD.

- Si vous n'utilisez pas l'ACL, ajoutez les autorisations suivantes :

Microsoft Graph	SharePoint
• Notes.Read.All (Application) : lit tous les blocs-notes OneNote • Sites.Read.All (Application) : lit les éléments de toutes les collections de sites	• AllSites.Read (Delegated) : lit les éléments de toutes les collections de sites

> **Note**
>
> Note.Read.All et Sites.Read.All sont obligatoires uniquement si vous souhaitez analyser des documents. OneNote
> Si vous souhaitez explorer des sites spécifiques, l'autorisation peut être limitée à des sites spécifiques plutôt qu'à tous les sites disponibles dans le domaine. Vous configurez l'autorisation Sites.Selected (Application). Avec cette autorisation d'API, vous devez définir explicitement l'autorisation d'accès sur chaque site via l'API Microsoft Graph. Pour plus d'informations, consultez le blog de Microsoft sur Sites.Autorisations sélectionnées.

- Si vous utilisez ACL, ajoutez les autorisations suivantes :

Microsoft Graph	SharePoint
• Group.Member.Read.All (Application) : lit toutes les adhésions aux groupes • Notes.Read.All (Application) : lit tous les blocs-notes OneNote • Des sites. FullControl.All (Delegated) : obligatoire pour récupérer ACLs les documents • Sites.Read.All (Application) : lit les éléments de toutes les collections de sites • User.Read.All (Application) : lit les profils complets de tous les utilisateurs	• AllSites.Read (Delegated) : lit les éléments de toutes les collections de sites

> **Note**
>
> GroupMember.Read.All et User.Read.All ne sont obligatoires que si Identity Crawler est activé.
> Si vous souhaitez explorer des sites spécifiques, l'autorisation peut être limitée à des sites spécifiques plutôt qu'à tous les sites disponibles dans le domaine. Vous configurez l'autorisation Sites.Selected (Application). Avec cette autorisation d'API, vous devez définir explicitement l'autorisation d'accès sur chaque site via l'API Microsoft Graph. Pour plus d'informations, consultez le blog de Microsoft sur Sites.Autorisations sélectionnées.

- Pour l'authentification réservée aux applications Azure AD : clé privée et ID client que vous avez générés après votre inscription SharePoint auprès d'Azure AD. Notez également le certificat X.509.
  - Si vous n'utilisez pas l'ACL, ajoutez les autorisations suivantes :

SharePoint
• Sites.Read.All (Application) : obligatoire pour accéder aux éléments et aux listes de toutes les collections de sites

> **Note**
>
> Si vous souhaitez explorer des sites spécifiques, l'autorisation peut être limitée à des sites spécifiques plutôt qu'à tous les sites disponibles dans le domaine. Vous configurez l'autorisation Sites.Selected (Application). Avec cette autorisation d'API, vous devez définir explicitement l'autorisation d'accès sur chaque site via l'API Microsoft Graph. Pour plus d'informations, consultez le blog de Microsoft sur Sites.Autorisations sélectionnées.

- Si vous utilisez ACL, ajoutez les autorisations suivantes :

SharePoint
• Des sites. FullControl.All (Application) — Nécessaire pour récupérer ACLs les documents

> **Note**
>
> Si vous souhaitez explorer des sites spécifiques, l'autorisation peut être limitée à des sites spécifiques plutôt qu'à tous les sites disponibles dans le domaine. Vous configurez l'autorisation Sites.Selected (Application). Avec cette autorisation d'API, vous devez définir explicitement l'autorisation d'accès sur chaque site via l'API Microsoft Graph. Pour plus d'informations, consultez le blog de Microsoft sur Sites.Autorisations sélectionnées.

- Pour l'authentification SharePoint uniquement par application : notez votre identifiant SharePoint client et votre code secret client générés lors de l'octroi de l'autorisation à SharePoint App Only, ainsi que votre identifiant client et votre secret client générés lorsque vous avez enregistré votre SharePoint application auprès d'Azure AD.

> **Note**
>
> SharePoint L'authentification par application uniquement n'est pas prise en charge pour la version SharePoint 2013.

- (Facultatif) Si vous analysez OneNote des documents et utilisez Identity Crawler, ajoutez les autorisations suivantes :

Microsoft Graph
• GroupMember.Read.All (Application) : lit toutes les adhésions aux groupes • Notes.Read.All (Application) : lit tous les blocs-notes OneNote • Sites.Read.All (Application) : lit les éléments de toutes les collections de sites • User.Read.All (Application) : lit les profils complets de tous les utilisateurs

> **Note**
>
> Aucune autorisation d'API n'est requise pour analyser des entités à l'aide de l'authentification de base et de l'authentification uniquement SharePoint par application.

Dans SharePoint Server, assurez-vous que vous disposez des éléments suivants :

- Vous avez copié votre SharePoint instance URLs et le nom de domaine de votre SharePoint URLs. Le format de l'URL d'hôte que vous entrez est*`https://yourcompany/sites/mysite`*. Votre URL doit commencer par`https`.

> **Note**
>
> (Sur place/sur serveur) Amazon Kendra vérifie si les informations de point de terminaison incluses sont les mêmes AWS Secrets Manager que celles spécifiées dans les détails de configuration de votre source de données. Cela permet de se protéger contre le problème de confusion des adjoints, qui est un problème de sécurité lorsqu'un utilisateur n'est pas autorisé à effectuer une action mais l'utilise Amazon Kendra comme proxy pour accéder au secret configuré et exécuter l'action. Si vous modifiez ultérieurement les informations

de votre point de terminaison, vous devez créer un nouveau secret pour synchroniser ces informations.

- L'authentification multifactorielle (MFA) a été désactivée dans votre SharePoint compte, afin de ne pas empêcher l' Amazon Kendra exploration de votre contenu. SharePoint
- Si vous utilisez l'authentification SharePoint App-Only pour le contrôle d'accès :
  - Vous avez copié l'ID SharePoint client généré lors de l'enregistrement de l'application uniquement au niveau du site. Le format de l'ID client est ClientId @TenantId. Par exemple, *ffa956f3-8f89-44e7-b0e4-49670756342c@888d0b57-69f1-4fb8-957f-e1f0bedf82fe*.
  - Vous avez copié le secret SharePoint client généré lors de l'enregistrement de l'application uniquement au niveau du site.

  Remarque : étant donné que les secrets client IDs et client sont générés pour des sites uniques uniquement lorsque vous enregistrez le SharePoint serveur pour l'authentification App Only, une seule URL de site est prise en charge pour SharePoint l'authentification App Only.

  > **Note**
  >
  > SharePoint L'authentification par application uniquement n'est pas prise en charge pour la version SharePoint 2013.

- Si vous utilisez un identifiant e-mail avec un domaine personnalisé pour le contrôle d'accès :
  - Vous avez noté la valeur de votre domaine de messagerie personnalisé, par exemple : « »*amazon.com*.
- Si vous utilisez l'adresse e-mail avec l'autorisation Domain from IDP, copiez votre :
  - Point de terminaison du serveur LDAP (point de terminaison du serveur LDAP, y compris le protocole et le numéro de port). Par exemple : *ldap://example.com:389*.
  - Base de recherche LDAP (base de recherche de l'utilisateur LDAP). Par exemple : *CN=Users,DC=sharepoint,DC=com*.
  - Nom d'utilisateur et mot de passe LDAP.
- Vous avez configuré des informations d'authentification NTLM ou des informations d'authentification Kerberos configurées contenant un nom d'utilisateur (nom d'utilisateur du SharePoint compte) et un mot de passe (mot de passe du compte). SharePoint

Dans votre Compte AWS, assurez-vous d'avoir :

- Création d'un Amazon Kendra index et, si vous utilisez l'API, notez l'ID de l'index.
- Vous avez créé un IAM rôle pour votre source de données et, si vous utilisez l'API, notez l'ARN du IAM rôle.

> **Note**
>
> Si vous modifiez votre type d'authentification et vos informations d'identification, vous devez mettre à jour votre IAM rôle pour accéder au bon identifiant AWS Secrets Manager secret.

- Stockez vos informations d' SharePoint authentification dans un AWS Secrets Manager secret et, si vous utilisez l'API, notez l'ARN du secret.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

Si vous n'avez pas de IAM rôle ou de secret existant, vous pouvez utiliser la console pour créer un nouveau IAM rôle et un nouveau Secrets Manager secret lorsque vous connectez votre source de SharePoint données à Amazon Kendra. Si vous utilisez l'API, vous devez fournir l'ARN d'un IAM rôle et d'un Secrets Manager secret existants, ainsi qu'un identifiant d'index.

Instructions de connexion

Pour vous connecter Amazon Kendra à votre source de SharePoint données, vous devez fournir les détails de vos SharePoint informations d'identification afin de Amazon Kendra pouvoir accéder à vos données. Si vous n'avez pas encore configuré SharePoint pour Amazon Kendra voirPrérequis.

Console: SharePoint Online

Pour vous connecter Amazon Kendra à SharePoint Online

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la.
2. Dans le volet de navigation de gauche, choisissez Index, puis choisissez l'index que vous souhaitez utiliser dans la liste des index.

> **Note**
>
> Vous pouvez choisir de configurer ou de modifier vos paramètres de contrôle d'accès utilisateur dans les paramètres de l'index.

3. Sur la page de démarrage, choisissez Ajouter une source de données.
4. Sur la page Ajouter une source de données, choisissez SharePoint connecteur, puis sélectionnez Ajouter un connecteur. Si vous utilisez la version 2 (le cas échéant), choisissez le SharePoint connecteur avec la balise « V2.0 ».
5. Sur la page Spécifier les détails de la source de données, entrez les informations suivantes :
   a. Dans Nom et description, pour Nom de la source de données : entrez le nom de votre source de données. Vous pouvez inclure des tirets, mais pas des espaces.
   b. (Facultatif) Description : entrez une description facultative pour votre source de données.
   c. Dans la langue par défaut : choisissez une langue pour filtrer vos documents pour l'index. Sauf indication contraire, la langue par défaut est l'anglais. La langue spécifiée dans les métadonnées du document remplace la langue sélectionnée.
   d. Dans Balises, pour Ajouter une nouvelle balise : incluez des balises facultatives pour rechercher et filtrer vos ressources ou suivre vos AWS coûts.
   e. Choisissez Suivant.
6. Sur la page Définir l'accès et la sécurité, entrez les informations suivantes :
   a. Méthode d'hébergement —Choisissez SharePointen ligne.
   b. Site URLs spécifique à votre SharePoint dépôt —Entrez l' SharePoint hôte URLs. Le format de l'hôte URLs que vous entrez est*`https://yourdomain.sharepoint.com/sites/mysite`*. L'URL doit commencer par le `https` protocole. Séparez URLs par une nouvelle ligne. Vous pouvez en ajouter jusqu'à 100 URLs.
   c. Domaine —Entrez le SharePoint domaine. Par exemple, le domaine de l'URL *`https://yourdomain.sharepoint.com/sites/mysite`* est*`yourdomain`*.
   d. Autorisation : activez ou désactivez les informations de la liste de contrôle d'accès (ACL) pour vos documents, si vous disposez d'une ACL et que vous souhaitez l'utiliser

pour le contrôle d'accès. L'ACL indique les documents auxquels les utilisateurs et les groupes peuvent accéder. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.

Vous pouvez également choisir le type d'ID utilisateur, qu'il s'agisse du nom d'utilisateur principal ou de l'e-mail de l'utilisateur extrait du portail Azure. Si vous ne le spécifiez pas, le courrier électronique est utilisé par défaut.

e. Authentification : choisissez l'authentification de base, OAuth 2.0, l'authentification Azure AD uniquement, l'authentification uniquement les SharePoint applications ou OAuth l'authentification par jeton d'actualisation 2.0. Vous pouvez soit choisir un AWS Secrets Manager secret existant pour stocker vos informations d'authentification, soit créer un secret.

   i. Si vous utilisez l'authentification de base, votre code secret doit inclure un nom secret, un nom SharePoint d'utilisateur et un mot de passe.

   ii. Si vous utilisez l'authentification OAuth 2.0, votre code secret doit inclure l'identifiant du SharePoint locataire, le nom secret, le nom SharePoint d'utilisateur, le mot de passe, l'identifiant du client Azure AD généré lors de votre inscription SharePoint dans Azure AD et le secret du client Azure AD généré lorsque vous vous inscrivez SharePoint dans Azure AD.

   iii. Si vous utilisez l'authentification Azure AD App-Only, votre code secret doit inclure l'ID du SharePoint locataire, le certificat X.509 autosigné Azure AD, le nom secret, l'identifiant du client Azure AD généré lors de votre inscription SharePoint dans Azure AD et la clé privée pour authentifier le connecteur pour Azure AD.

   iv. Si vous utilisez l'authentification SharePoint App-Only, votre code secret doit inclure l' SharePoint identifiant du client, le nom secret, l'identifiant SharePoint client que vous avez généré lors de l'enregistrement de l'application uniquement au niveau du locataire, le secret SharePoint client généré lors de votre inscription à App Only au niveau du locataire, l'identifiant client Azure AD généré lorsque vous vous inscrivez SharePoint dans Azure AD et le secret client Azure AD généré lorsque vous vous inscrivez SharePoint sur Azure AD.

   Le format de l'ID SharePoint client est*ClientID@TenantId*. Par exemple, *ffa956f3-8f89-44e7-b0e4-49670756342c@888d0b57-69f1-4fb8-957f-e1f0bedf82fe*.

v. Si vous utilisez l'authentification par jeton d'actualisation OAuth 2.0, votre code secret doit inclure l'identifiant du SharePoint locataire, le nom du secret, l'identifiant client Azure AD unique généré lorsque vous vous inscrivez SharePoint dans Azure AD, le secret du client Azure AD généré lorsque vous vous inscrivez SharePoint sur Azure AD, le jeton d'actualisation généré pour vous connecter Amazon Kendra à SharePoint.

f. Virtual Private Cloud (VPC) —Vous pouvez choisir d'utiliser un VPC. Dans ce cas, vous devez ajouter des sous-réseaux et des groupes de sécurité VPC.

g. Identity Crawler : spécifiez s'il faut activer l'explorateur Amazon Kendra d'identité. Le robot d'exploration d'identité utilise les informations de la liste de contrôle d'accès (ACL) de vos documents pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Si vous disposez d'une ACL pour vos documents et que vous choisissez de l'utiliser, vous pouvez également choisir d'activer le robot d'exploration Amazon Kendra d'identité pour configurer le filtrage des résultats de recherche par contexte utilisateur. Sinon, si le robot d'identification est désactivé, tous les documents peuvent être consultés publiquement. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'exploration d'identité est désactivé, vous pouvez également utiliser l'PutPrincipalMappingAPI pour télécharger les informations d'accès des utilisateurs et des groupes afin de filtrer le contexte utilisateur.

   Vous pouvez également choisir d'explorer le mappage de groupe local ou le mappage de groupe Azure Active Directory.

   > **Note**
   >
   > L'analyse cartographique du groupe AD n'est disponible que pour les authentifications OAuth OAuth 2.0, 2.0 et SharePoint App Only.

h. IAM rôle —Choisissez un IAM rôle existant ou créez-en un nouveau IAM pour accéder aux informations d'identification de votre référentiel et indexer le contenu.

   > **Note**
   >
   > IAM les rôles utilisés pour les index ne peuvent pas être utilisés pour les sources de données. Si vous ne savez pas si un rôle existant est utilisé pour un index ou une FAQ, choisissez Créer un nouveau rôle pour éviter les erreurs.

i. Choisissez Suivant.

7. Sur la page Configurer les paramètres de synchronisation, entrez les informations suivantes :

   a. Dans le champ d'application de la synchronisation, choisissez l'une des options suivantes :

      i. Sélectionnez les entités : choisissez les entités que vous souhaitez analyser. Vous pouvez choisir d'explorer toutes les entités ou toute combinaison de fichiers, de pièces jointes, de pages de liens, d'événements, de commentaires et de données de liste.

      ii. Dans Configuration supplémentaire, pour les modèles d'expression régulière d'entité : ajoutez des modèles d'expressions régulières pour les liens, les pages et les événements afin d'inclure des entités spécifiques au lieu de synchroniser tous vos documents.

      iii. Modèles Regex : ajoutez des modèles d'expressions régulières pour inclure ou exclure des fichiers par chemin de fichier, nom de fichier, type de fichier, nom de OneNote section et nom de OneNote page au lieu de synchroniser tous vos documents. Vous pouvez en ajouter jusqu'à 100.

      > **Note**
      >
      > OneNote l'exploration n'est disponible que pour la OAuth version OAuth 2.0, le jeton d'actualisation 2.0 et SharePoint l'authentification App Only.

   b. Pour le mode de synchronisation, choisissez la manière dont vous souhaitez mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est synchronisé par défaut.

      - Synchronisation complète : synchronisez tout le contenu, quel que soit l'état de synchronisation précédent.
      - Synchronisation des documents nouveaux ou modifiés : synchronisez uniquement les documents nouveaux ou modifiés.
      - Synchronisation des documents nouveaux, modifiés ou supprimés : synchronisez uniquement les documents nouveaux, modifiés et supprimés.

c. Dans Calendrier d'exécution de la synchronisation, pour Fréquence : choisissez la fréquence à laquelle vous souhaitez synchroniser le contenu de votre source de données et mettre à jour votre index.

d. Choisissez Suivant.

8. Sur la page Définir les mappages de champs, entrez les informations suivantes :

a. Champs de source de données par défaut : sélectionnez parmi les champs de source de données par défaut Amazon Kendra générés que vous souhaitez mapper à votre index.

b. Ajouter un champ —Pour ajouter des champs de source de données personnalisés afin de créer un nom de champ d'index à mapper et le type de données du champ.

c. Choisissez Suivant.

9. Sur la page Réviser et créer, vérifiez que les informations saisies sont correctes, puis sélectionnez Ajouter une source de données. Vous pouvez également choisir de modifier vos informations à partir de cette page. Votre source de données apparaîtra sur la page Sources de données une fois que la source de données aura été ajoutée avec succès.

Console: SharePoint Server

Pour vous connecter Amazon Kendra à SharePoint

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la.
2. Dans le volet de navigation de gauche, choisissez Index, puis choisissez l'index que vous souhaitez utiliser dans la liste des index.

> **Note**
>
> Vous pouvez choisir de configurer ou de modifier vos paramètres de contrôle d'accès utilisateur dans les paramètres de l'index.

3. Sur la page de démarrage, choisissez Ajouter une source de données.
4. Sur la page Ajouter une source de données, choisissez SharePoint connecteur, puis sélectionnez Ajouter un connecteur. Si vous utilisez la version 2 (le cas échéant), choisissez le SharePoint connecteur avec la balise « V2.0 ».
5. Sur la page Spécifier les détails de la source de données, entrez les informations suivantes :

a. Dans Nom et description, pour Nom de la source de données : entrez le nom de votre source de données. Vous pouvez inclure des tirets, mais pas des espaces.

b. (Facultatif) Description : entrez une description facultative pour votre source de données.

c. Dans la langue par défaut : choisissez une langue pour filtrer vos documents pour l'index. Sauf indication contraire, la langue par défaut est l'anglais. La langue spécifiée dans les métadonnées du document remplace la langue sélectionnée.

d. Dans Balises, pour Ajouter une nouvelle balise : incluez des balises facultatives pour rechercher et filtrer vos ressources ou suivre vos AWS coûts.

e. Choisissez Suivant.

6. Sur la page Définir l'accès et la sécurité, entrez les informations suivantes :

a. Méthode d'hébergement : choisissez le SharePointserveur.

b. Choisissez SharePoint la version —Choisissez SharePoint 2013, SharePoint 2016, SharePoint 2019 et SharePoint (édition par abonnement).

c. Site URLs spécifique à votre SharePoint dépôt —Entrez l' SharePoint hôte URLs. Le format de l'hôte URLs que vous entrez est*`https://yourcompany/sites/mysite`*. L'URL doit commencer par le `https` protocole. Séparez URLs par une nouvelle ligne. Vous pouvez en ajouter jusqu'à 100 URLs.

d. Domaine —Entrez le SharePoint domaine. Par exemple, le domaine de l'URL *`https://yourcompany/sites/mysite`* est *`yourcompany`*

e. Emplacement du certificat SSL —Entrez le Amazon S3 chemin d'accès à votre fichier de certificat SSL.

f. (Facultatif) Pour le proxy Web : entrez le nom d'hôte (sans le `https://` protocole `http://` OR) et le numéro de port utilisé par le protocole de transport d'URL de l'hôte. La valeur numérique du numéro de port doit être comprise entre 0 et 65535.

g. Autorisation : activez ou désactivez les informations de la liste de contrôle d'accès (ACL) pour vos documents, si vous disposez d'une ACL et que vous souhaitez l'utiliser pour le contrôle d'accès. L'ACL indique les documents auxquels les utilisateurs et les groupes peuvent accéder. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.

Pour le SharePoint serveur, vous pouvez choisir l'une des options ACL suivantes :

i. ID e-mail avec domaine provenant de l'IDP : l'ID utilisateur est basé sur le courrier électronique dont IDs les domaines ont été récupérés auprès du fournisseur d'identité (IDP) sous-jacent. Vous fournissez les détails de connexion IDP dans votre Secrets Manager secret dans le cadre de l'authentification.

ii. ID e-mail avec domaine personnalisé : l'ID utilisateur est basé sur la valeur du domaine de messagerie personnalisé. Par exemple, « *amazon.com* ». Le domaine de messagerie sera utilisé pour créer l'identifiant de courrier électronique pour le contrôle d'accès. Vous devez saisir votre domaine de messagerie personnalisé.

iii. Domaine \ Utilisateur avec domaine : l'ID utilisateur est construit au format Domain \ User ID. Vous devez fournir un nom de domaine valide. Par exemple : pour *"sharepoint2019"* créer un contrôle d'accès.

h. Pour Authentification, choisissez l'authentification SharePoint App-Only, l'authentification NTLM ou l'authentification Kerberos. Vous pouvez soit choisir un AWS Secrets Manager secret existant pour stocker vos informations d'authentification, soit créer un secret.

i. Si vous utilisez l'authentification NTLM ou Kerberos, votre code secret doit inclure un nom secret, un nom d' SharePoint utilisateur et un mot de passe.

   Si vous utilisez l'adresse e-mail avec le domaine de l'IDP, entrez également votre :

   - Point de terminaison du serveur LDAP : point de terminaison du serveur LDAP, y compris le protocole et le numéro de port. Par exemple : *ldap://example.com:389*.
   - Base de recherche LDAP —Base de recherche de l'utilisateur LDAP. Par exemple : *CN=Users,DC=sharepoint,DC=com*.
   - Nom d'utilisateur LDAP : votre nom d'utilisateur LDAP.
   - Mot de passe LDAP : votre mot de passe LDAP.

ii. Si vous utilisez l'authentification SharePoint App Only, votre code secret doit inclure un nom secret, l'ID SharePoint client que vous avez généré lorsque vous avez enregistré App Only au niveau du site, le secret SharePoint client généré lorsque vous vous inscrivez à App Only at Site Level.

   Le format de l'ID SharePoint client est*ClientID@TenantId*. Par exemple, *ffa956f3-8f89-44e7-b0e4-49670756342c@888d0b57-69f1-4fb8-957f-e1f0bedf82fe*.

Remarque : étant donné que les secrets client IDs et client sont générés pour des sites uniques uniquement lorsque vous enregistrez le SharePoint serveur pour l'authentification App Only, une seule URL de site est prise en charge pour SharePoint l'authentification App Only.

Si vous utilisez l'adresse e-mail avec le domaine de l'IDP, entrez également votre :

- Point de terminaison du serveur LDAP : point de terminaison du serveur LDAP, y compris le protocole et le numéro de port. Par exemple : *`ldap://example.com:389`*.
- Base de recherche LDAP —Base de recherche de l'utilisateur LDAP. Par exemple : *`CN=Users,DC=sharepoint,DC=com`*.
- Nom d'utilisateur LDAP : votre nom d'utilisateur LDAP.
- Mot de passe LDAP : votre mot de passe LDAP.

i. Virtual Private Cloud (VPC) —Vous pouvez choisir d'utiliser un VPC. Dans ce cas, vous devez ajouter des sous-réseaux et des groupes de sécurité VPC.

j. Identity Crawler : spécifiez s'il faut activer l'explorateur Amazon Kendra d'identité. Le robot d'exploration d'identité utilise les informations de la liste de contrôle d'accès (ACL) de vos documents pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Si vous disposez d'une ACL pour vos documents et que vous choisissez de l'utiliser, vous pouvez également choisir d'activer le robot d'exploration Amazon Kendra d'identité pour configurer le filtrage des résultats de recherche par contexte utilisateur. Sinon, si le robot d'identification est désactivé, tous les documents peuvent être consultés publiquement. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'exploration d'identité est désactivé, vous pouvez également utiliser l'PutPrincipalMappingAPI pour télécharger les informations d'accès des utilisateurs et des groupes afin de filtrer le contexte utilisateur.

   Vous pouvez également choisir d'explorer le mappage de groupe local ou le mappage de groupe Azure Active Directory.

   > **Note**
   >
   > L'analyse cartographique du groupe AD est disponible uniquement avec l'authentification SharePoint App Only.

k. IAM rôle —Choisissez un IAM rôle existant ou créez-en un nouveau IAM pour accéder aux informations d'identification de votre référentiel et indexer le contenu.

> **Note**
>
> IAM les rôles utilisés pour les index ne peuvent pas être utilisés pour les sources de données. Si vous ne savez pas si un rôle existant est utilisé pour un index ou une FAQ, choisissez Créer un nouveau rôle pour éviter les erreurs.

l. Choisissez Suivant.

7. Sur la page Configurer les paramètres de synchronisation, entrez les informations suivantes :

a. Dans le champ d'application de la synchronisation, choisissez l'une des options suivantes :

i. Sélectionnez les entités : choisissez les entités que vous souhaitez analyser. Vous pouvez choisir d'explorer toutes les entités ou toute combinaison de fichiers, de pièces jointes, de pages de liens, d'événements et de données de liste.

ii. Dans Configuration supplémentaire, pour les modèles d'expression régulière d'entité : ajoutez des modèles d'expressions régulières pour les liens, les pages et les événements afin d'inclure des entités spécifiques au lieu de synchroniser tous vos documents.

iii. Modèles Regex : ajoutez des modèles d'expressions régulières pour inclure ou exclure des fichiers par chemin de fichier Nom de fichier Type de fichier, nom de OneNote section et nom de OneNote page au lieu de synchroniser tous vos documents. Vous pouvez en ajouter jusqu'à 100.

> **Note**
>
> OneNote l'exploration n'est disponible que pour SharePoint l'authentification par application uniquement.

b. Mode de synchronisation : choisissez la manière dont vous souhaitez mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation

complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation.

- Synchronisation complète : Indexez tout le contenu fraîchement, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
- Nouvelle synchronisation modifiée : Indexez uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
- Synchronisation nouvelle, modifiée ou supprimée : Indexez uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

c. Dans Calendrier d'exécution de la synchronisation, pour Fréquence : choisissez la fréquence à laquelle vous souhaitez synchroniser le contenu de votre source de données et mettre à jour votre index.

d. Choisissez Suivant.

8. Sur la page Définir les mappages de champs, entrez les informations suivantes :

a. Champs de source de données par défaut : sélectionnez parmi les champs de source de données par défaut Amazon Kendra générés que vous souhaitez mapper à votre index.

b. Ajouter un champ —Pour ajouter des champs de source de données personnalisés afin de créer un nom de champ d'index à mapper et le type de données du champ.

c. Choisissez Suivant.

9. Sur la page Réviser et créer, vérifiez que les informations saisies sont correctes, puis sélectionnez Ajouter une source de données. Vous pouvez également choisir de modifier vos informations à partir de cette page. Votre source de données apparaîtra sur la page Sources de données une fois que la source de données aura été ajoutée avec succès.

API

Pour vous connecter Amazon Kendra à SharePoint

Vous devez spécifier un JSON du schéma de source de données à l'aide de l'TemplateConfigurationAPI. Vous devez fournir les informations suivantes :

- Source de données —Spécifiez le type de source de données tel que `SHAREPOINTV2` lorsque vous utilisez le schéma TemplateConfigurationJSON. Spécifiez également la source de données `TEMPLATE` lorsque vous appelez l'CreateDataSourceAPI.
- Métadonnées du point de terminaison du référentiel : spécifiez le `tenantID` `domain` et `siteUrls` de votre SharePoint instance.
- Mode de synchronisation : spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation. Vous pouvez choisir entre :
  - `FORCED_FULL_CRAWL`pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
  - `FULL_CRAWL`pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
  - `CHANGE_LOG`pour indexer uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
- Identity Crawler : spécifiez s'il faut activer l'explorateur Amazon Kendra d'identité. Le robot d'exploration d'identité utilise les informations de la liste de contrôle d'accès (ACL) de vos documents pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Si vous disposez d'une ACL pour vos documents et que vous choisissez de l'utiliser, vous pouvez également choisir d'activer le robot d'exploration Amazon Kendra d'identité pour configurer le filtrage des résultats de recherche par contexte utilisateur. Sinon, si le robot d'identification est désactivé, tous les documents peuvent être consultés publiquement. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'exploration d'identité est désactivé, vous pouvez également utiliser l'PutPrincipalMappingAPI pour télécharger les informations d'accès des utilisateurs et des groupes afin de filtrer le contexte utilisateur.

> **Note**
>
> Identity Crawler n'est disponible que lorsque vous le configurez sur`crawlAcl`. `true`

- Propriétés supplémentaires du référentiel : spécifiez les éléments suivants :
  - (pour Azure AD) `s3bucketName` et `s3certificateName` que vous utilisez pour stocker votre certificat X.509 autosigné Azure AD.
  - Type d'authentification (`auth_Type`) que vous utilisez`OAuth2`, que ce soit `OAuth2App` `OAuth2CertificateBasic`,,`OAuth2_RefreshToken`,`NTLM`, et`Kerberos`.
  - Version (`version`) que vous utilisez, que ce soit `Server` ou`Online`. Si vous en utilisez, `Server` vous pouvez spécifier davantage le `onPremVersion` as`2013`, `20162019`, ou`SubscriptionEdition`.
- Nom de ressource Amazon (ARN) secret : indiquez le nom de ressource Amazon (ARN) d'un Secrets Manager secret contenant les informations d'authentification que vous avez créées dans votre SharePoint compte.

  Si vous utilisez SharePoint Online, vous pouvez choisir entre l'authentification de base, l'authentification OAuth 2.0, l'authentification Azure AD uniquement et SharePoint l'authentification uniquement pour les applications. Voici la structure JSON minimale qui doit figurer dans votre code secret pour chaque option d'authentification :

  - Authentification de base

    ```
    {
        "userName": "SharePoint account user name",
        "password": "SharePoint account password"
    }
    ```

  - OAuth Authentification 2.0

    ```
    {
        "clientId": "client id generated when registering SharePoint with Azure AD",
        "clientSecret": "client secret generated when registering SharePoint with
     Azure AD",
        "userName": "SharePoint account user name",
        "password": "SharePoint account password"
    }
    ```

  - Authentification uniquement avec les applications Azure AD

```
{
    "clientId": "client id generated when registering SharePoint with Azure AD",
    "privateKey": "private key to authorize connection with Azure AD"
}
```

- SharePoint Authentification uniquement par application

```
{
    "clientId": "client id generated when registering SharePoint for App Only at
 Tenant Level",
    "clientSecret": "client secret generated when registering SharePoint for App
 Only at Tenant Level",
    "adClientId": "client id generated while registering SharePoint with Azure
 AD",
    "adClientSecret": "client secret generated while registering SharePoint with
 Azure AD"
}
```

- OAuth Authentification par jeton d'actualisation 2.0

```
{
    "clientId": "client id generated when registering SharePoint with Azure AD",
    "clientSecret": "client secret generated when registering SharePoint with
 Azure AD",
    "refreshToken": "refresh token generated to connect to SharePoint"
}
```

Si vous utilisez SharePoint Server, vous pouvez choisir entre l'authentification SharePoint App-Only, l'authentification NTLM et l'authentification Kerberos. Voici la structure JSON minimale qui doit figurer dans votre code secret pour chaque option d'authentification :

- SharePoint Authentification uniquement par application

```
{
    "siteUrlsHash": "Hash representation of SharePoint site URLs",
    "clientId": "client id generated when registering SharePoint for App Only at
 Site Level",
    "clientSecret": "client secret generated when registering SharePoint for App
 Only at Site Level"
}
```

- SharePoint Authentification uniquement par application avec autorisation du domaine depuis l'IDP

```
{
    "siteUrlsHash": "Hash representation of SharePoint site URLs",
    "clientId": "client id generated when registering SharePoint for App Only at
 Site Level",
    "clientSecret": "client secret generated when registering SharePoint for App
 Only at Site Level",
    "ldapUrl": "LDAP Account url eg. ldap://example.com:389",
    "baseDn": "LDAP Account base dn eg. CN=Users,DC=sharepoint,DC=com",
    "ldapUser": "LDAP account user name",
    "ldapPassword": "LDAP account password"
}
```

- Authentification NTLM ou Kerberos (serveur uniquement)

```
{
    "siteUrlsHash": "Hash representation of SharePoint site URLs",
    "userName": "SharePoint account user name",
    "password": "SharePoint account password"
}
```

- Authentification NTLM ou Kerberos (serveur uniquement) avec autorisation de domaine depuis l'IDP

```
{
    "siteUrlsHash": "Hash representation of SharePoint site URLs",
    "userName": "SharePoint account user name",
    "password": "SharePoint account password",
    "ldapUrl": "ldap://example.com:389",
    "baseDn": "CN=Users,DC=sharepoint,DC=com",
    "ldapUser": "LDAP account user name",
    "ldapPassword": "LDAP account password"
}
```

- IAM role —Spécifiez le `RoleArn` moment où vous appelez `CreateDataSource` pour fournir à un IAM rôle les autorisations d'accéder à votre Secrets Manager secret et d'appeler le public requis APIs pour le SharePoint connecteur et Amazon Kendra. Pour plus d'informations, consultez la section IAM Rôles des sources de SharePoint données.

Vous pouvez également ajouter les fonctionnalités optionnelles suivantes :

- Virtual Private Cloud (VPC) **`VpcConfiguration`** —Spécifiez le moment de votre appel. `CreateDataSource` Pour de plus amples informations, veuillez consulter Configuration Amazon Kendra pour utiliser un Amazon VPC.
- Filtres d'inclusion et d'exclusion : vous pouvez spécifier si vous souhaitez inclure ou exclure certains fichiers et autres contenus. OneNotes

> **Note**
>
> La plupart des sources de données utilisent des modèles d'expressions régulières, qui sont des modèles d'inclusion ou d'exclusion appelés filtres. Si vous spécifiez un filtre d'inclusion, seul le contenu correspondant au filtre d'inclusion est indexé. Tout document qui ne correspond pas au filtre d'inclusion n'est pas indexé. Si vous spécifiez un filtre d'inclusion et d'exclusion, les documents correspondant au filtre d'exclusion ne sont pas indexés, même s'ils correspondent au filtre d'inclusion.

- Mappages de champs : choisissez de mapper les champs de votre source de SharePoint données à vos champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.

> **Note**
>
> Le champ du corps du document ou l'équivalent du corps du document pour vos documents est requis pour Amazon Kendra effectuer une recherche dans vos documents. Vous devez associer le nom du champ du corps du document dans votre source de données au nom du champ d'index`_document_body`. Tous les autres champs sont facultatifs.

Pour obtenir la liste des autres clés JSON importantes à configurer, consultez le schéma du SharePoint modèle.

Remarques

- Le connecteur prend en charge les mappages de champs personnalisés uniquement pour l'entité Files.

- Pour toutes les versions SharePoint du serveur, le jeton ACL doit être en minuscules. Pour un e-mail avec un domaine provenant de l'IDP et un identifiant de messagerie avec une ACL de domaine personnalisée, par exemple :*user@sharepoint2019.com*. Pour Domain \ User with Domain ACL, par exemple :*sharepoint2013\user*.
- Lorsque les listes de contrôle d'accès (ACLs) sont activées, l'option « Synchroniser uniquement le contenu nouveau ou modifié » n'est pas disponible en raison des limites de SharePoint l'API. Nous vous recommandons d'utiliser plutôt les modes « Synchronisation complète » ou « Synchronisation du contenu nouveau, modifié ou supprimé », ou de les désactiver ACLs si vous devez utiliser ce mode de synchronisation.
- Le connecteur ne prend pas en charge le mode journal des changements/la synchronisation du contenu nouveau ou modifié pour SharePoint 2013.
- Si le nom d'une entité contient un % caractère, le connecteur ignorera ces fichiers en raison des limites de l'API.
- OneNote ne peut être exploré par le connecteur qu'à l'aide d'un identifiant de locataire et avec un jeton d'actualisation OAuth OAuth 2.0, 2.0 ou une authentification SharePoint App Only activée pour SharePoint Online.
- Le connecteur explore la première section d'un OneNote document en utilisant uniquement son nom par défaut, même si le document est renommé.
- Le connecteur analyse les liens dans les éditions SharePoint 2019, SharePoint en ligne et par abonnement, uniquement si les pages et les fichiers sont sélectionnés comme entités à analyser en plus des liens.
- Le connecteur analyse les liens en SharePoint 2013 et SharePoint 2016 si Links est sélectionné comme entité à analyser.
- Le connecteur analyse les pièces jointes et les commentaires des listes uniquement lorsque List Data est également sélectionnée comme entité à analyser.
- Le connecteur analyse les pièces jointes aux événements uniquement lorsque Events est également sélectionné comme entité à analyser.
- Pour la version SharePoint en ligne, le jeton ACL sera en minuscules. Par exemple, si le nom principal de l'utilisateur se trouve *MaryMajor@domain.com* dans le portail Azure, le jeton ACL du SharePoint connecteur sera*marymajor@domain.com*.
- Dans Identity Crawler for SharePoint Online and Server, si vous souhaitez explorer des groupes imbriqués, vous devez activer Local et AD Group Crawling.
- Si vous utilisez SharePoint Online et que le nom d'utilisateur principal de votre portail Azure est une combinaison de majuscules et de minuscules, l' SharePoint API le convertit en interne

en minuscules. Pour cette raison, le Amazon Kendra SharePoint connecteur définit l'ACL en minuscules.

Microsoft SQL Server

> **Note**
>
> Microsoft SQL ServerLe connecteur reste entièrement pris en charge pour les clients existants jusqu'au 31 mai 2026. Bien que ce connecteur ne soit plus disponible pour les nouveaux utilisateurs, les utilisateurs actuels peuvent continuer à l'utiliser sans interruption. Nous faisons constamment évoluer notre portefeuille de connecteurs afin de proposer des solutions plus évolutives et personnalisables. Pour les intégrations futures, nous vous recommandons d'explorer le framework de connecteurs personnalisés Amazon Kendra [1], conçu pour prendre en charge un plus large éventail de cas d'utilisation en entreprise avec une flexibilité accrue.

Microsoft SQL Serverest un système de gestion de base de données relationnelle (RDBMS) développé par Microsoft. Si vous êtes un Microsoft SQL Server utilisateur, vous pouvez l'utiliser Amazon Kendra pour indexer votre source de Microsoft SQL Server données. Le connecteur Amazon Kendra Microsoft SQL Server de source de données prend en charge MS SQL Server 2019.

Vous pouvez vous connecter Amazon Kendra à votre source de Microsoft SQL Server données à l'aide de la Amazon Kendra console et de l'TemplateConfigurationAPI.

Pour résoudre les problèmes liés à votre connecteur de source de Amazon Kendra Microsoft SQL Server données, consultezDépannage des sources de données.

Rubriques

- Fonctionnalités prises en charge
- Prérequis
- Instructions de connexion
- Remarques

Fonctionnalités prises en charge

- Mappages de champs

- Filtrage du contexte utilisateur
- Filtres d'inclusion/exclusion
- Synchronisation complète et incrémentielle du contenu
- Cloud privé virtuel (VPC)

Prérequis

Avant de pouvoir Amazon Kendra indexer votre source de Microsoft SQL Server données, apportez ces modifications à vos AWS comptes Microsoft SQL Server and.

DansMicrosoft SQL Server, assurez-vous d'avoir :

- Notez le nom d'utilisateur et le mot de passe de votre base de données.

> **Important**
>
> Il est recommandé de fournir des informations d'identification Amazon Kendra de base de données en lecture seule.

- Vous avez copié l'URL, le port et l'instance de votre hôte de base de données.
- Il est vérifié que chaque document est unique dans Microsoft SQL Server et entre les autres sources de données que vous prévoyez d'utiliser pour le même index. Chaque source de données que vous souhaitez utiliser pour un index ne doit pas contenir le même document dans toutes les sources de données. IDs Les documents sont globaux par rapport à un index et doivent être uniques par index.

Dans votre Compte AWS, assurez-vous d'avoir :

- Création d'un Amazon Kendra index et, si vous utilisez l'API, notez l'ID de l'index.
- Vous avez créé un IAM rôle pour votre source de données et, si vous utilisez l'API, notez l'ARN du IAM rôle.

> **Note**
>
> Si vous modifiez votre type d'authentification et vos informations d'identification, vous devez mettre à jour votre IAM rôle pour accéder au bon identifiant AWS Secrets Manager secret.

- Stockez vos informations d'Microsoft SQL Serverauthentification dans un AWS Secrets Manager secret et, si vous utilisez l'API, notez l'ARN du secret.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

Si vous n'avez pas de IAM rôle ou de secret existant, vous pouvez utiliser la console pour créer un nouveau IAM rôle et un nouveau Secrets Manager secret lorsque vous connectez votre source de Microsoft SQL Server données à Amazon Kendra. Si vous utilisez l'API, vous devez fournir l'ARN d'un IAM rôle et d'un Secrets Manager secret existants, ainsi qu'un identifiant d'index.

Instructions de connexion

Pour vous connecter Amazon Kendra à votre source de Microsoft SQL Server données, vous devez fournir les détails de vos Microsoft SQL Server informations d'identification afin de Amazon Kendra pouvoir accéder à vos données. Si vous n'avez pas encore configuré Microsoft SQL Server pour Amazon Kendra voirPrérequis.

Console

Pour vous connecter Amazon Kendra à Microsoft SQL Server

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la.
2. Dans le volet de navigation de gauche, choisissez Index, puis choisissez l'index que vous souhaitez utiliser dans la liste des index.

> **Note**
>
> Vous pouvez choisir de configurer ou de modifier vos paramètres de contrôle d'accès utilisateur dans les paramètres de l'index.

3. Sur la page de démarrage, choisissez Ajouter une source de données.
4. Sur la page Ajouter une source de données, choisissez Microsoft SQL Serverconnecteur, puis sélectionnez Ajouter un connecteur. Si vous utilisez la version 2 (le cas échéant), choisissez le Microsoft SQL Serverconnecteur avec la balise « V2.0 ».
5. Sur la page Spécifier les détails de la source de données, entrez les informations suivantes :
   a. Dans Nom et description, pour Nom de la source de données : entrez le nom de votre source de données. Vous pouvez inclure des traits d'union, mais pas des espaces.
   b. (Facultatif) Description : entrez une description facultative pour votre source de données.
   c. Dans la langue par défaut : choisissez une langue pour filtrer vos documents pour l'index. Sauf indication contraire, la langue par défaut est l'anglais. La langue spécifiée dans les métadonnées du document remplace la langue sélectionnée.
   d. Dans Balises, pour Ajouter une nouvelle balise : incluez des balises facultatives pour rechercher et filtrer vos ressources ou suivre vos AWS coûts.
   e. Choisissez Suivant.
6. Sur la page Définir l'accès et la sécurité, entrez les informations suivantes :
   a. Dans Source, entrez les informations suivantes :
   b. Hôte : entrez le nom d'hôte de la base de données.
   c. Port — Entrez le port de base de données.
   d. Instance — Entrez l'instance de base de données.
   e. Activer l'emplacement du certificat SSL : choisissez d'entrer le Amazon S3 chemin d'accès à votre fichier de certificat SSL.
   f. Dans Authentification, entrez les informations suivantes :
      - AWS Secrets Manager secret —Choisissez un secret existant ou créez-en un nouveau Secrets Manager pour stocker vos informations Microsoft SQL Server d'authentification. Si vous choisissez de créer un nouveau secret, une fenêtre AWS Secrets Manager secrète s'ouvre.

A. Entrez les informations suivantes dans la fenêtre Créer un AWS Secrets Manager secret :

I. Nom secret : le nom de votre secret. Le préfixe « AmazonKendra - Microsoft SQL Server - » est automatiquement ajouté à votre nom secret.

II. Pour le nom d'utilisateur et le mot de passe de la base de données, entrez les valeurs d'identification d'authentification que vous avez copiées depuis votre base de données.

B. Choisissez Enregistrer.

g. Virtual Private Cloud (VPC) —Vous pouvez choisir d'utiliser un VPC. Dans ce cas, vous devez ajouter des sous-réseaux et des groupes de sécurité VPC.

h. IAM rôle —Choisissez un IAM rôle existant ou créez-en un nouveau IAM pour accéder aux informations d'identification de votre référentiel et indexer le contenu.

> **Note**
>
> IAM les rôles utilisés pour les index ne peuvent pas être utilisés pour les sources de données. Si vous ne savez pas si un rôle existant est utilisé pour un index ou une FAQ, choisissez Créer un nouveau rôle pour éviter les erreurs.

i. Choisissez Suivant.

7. Sur la page Configurer les paramètres de synchronisation, entrez les informations suivantes :

a. Dans le champ d'application de la synchronisation, choisissez l'une des options suivantes :

- Requête SQL —Entrez des instructions de requête SQL telles que les opérations SELECT et JOIN. La taille des requêtes SQL doit être inférieure à 32 Ko. Amazon Kendra analysera tout le contenu de la base de données correspondant à votre requête.

> **Note**
>
> Si le nom d'une table contient des caractères spéciaux (non alphanumériques), vous devez utiliser des crochets autour du nom de la table. Par exemple, `select * from [my-database-table]`

- Colonne clé primaire : indiquez la clé primaire pour la table de base de données. Cela identifie une table au sein de votre base de données.
- Colonne de titre : indiquez le nom de la colonne de titre du document dans votre table de base de données.
- Colonne du corps : indiquez le nom de la colonne du corps du document dans votre table de base de données.

b. Dans Configuration supplémentaire — facultatif, choisissez l'une des options suivantes pour synchroniser un contenu spécifique au lieu de synchroniser tous les fichiers :

- Colonnes détectant les modifications : entrez le nom des colonnes qui Amazon Kendra seront utilisées pour détecter les modifications de contenu. Amazon Kendra réindexera le contenu en cas de modification de l'une de ces colonnes.
- IDsColonne utilisateur : entrez le nom de la colonne contenant l'utilisateur devant être autorisé IDs à accéder au contenu.
- Colonne Groupes : entrez le nom de la colonne contenant les groupes autorisés à accéder au contenu.
- URLsColonne source —Entrez le nom de la colonne contenant la source URLs à indexer.
- Colonne horodatage : entrez le nom de la colonne contenant les horodatages. Amazon Kendra utilise les informations d'horodatage pour détecter les modifications apportées à votre contenu et synchroniser uniquement le contenu modifié.
- Colonne fuseaux horaires : entrez le nom de la colonne contenant les fuseaux horaires du contenu à analyser.
- Format d'horodatage : entrez le nom de la colonne contenant les formats d'horodatage à utiliser pour détecter les modifications de contenu et resynchroniser votre contenu.

c. Mode de synchronisation : choisissez la manière dont vous souhaitez mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation.

- Synchronisation complète : Indexez tout le contenu fraîchement, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.

- Nouvelle synchronisation modifiée : Indexez uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
- Synchronisation nouvelle, modifiée ou supprimée : Indexez uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

d. Dans Synchroniser le calendrier d'exécution, pour Fréquence : à quelle fréquence Amazon Kendra sera synchronisée avec votre source de données.

e. Choisissez Suivant.

8. Sur la page Définir les mappages de champs, entrez les informations suivantes :

   a. Sélectionnez parmi les champs de source de données par défaut générés (Document IDs, Titres des documents et Source URLs) que vous souhaitez Amazon Kendra mapper à l'index.

   b. Ajouter un champ —Pour ajouter des champs de source de données personnalisés afin de créer un nom de champ d'index à mapper et le type de données du champ.

   c. Choisissez Suivant.

9. Sur la page Réviser et créer, vérifiez que les informations saisies sont correctes, puis sélectionnez Ajouter une source de données. Vous pouvez également choisir de modifier vos informations à partir de cette page. Votre source de données apparaîtra sur la page Sources de données une fois que la source de données aura été ajoutée avec succès.

API

Pour vous connecter Amazon Kendra à Microsoft SQL Server

Vous devez spécifier les éléments suivants à l'aide de l'TemplateConfigurationAPI :

- Source de données —Spécifiez le type de source de données tel que `JDBC` lorsque vous utilisez le schéma TemplateConfigurationJSON. Spécifiez également la source de données `TEMPLATE` lorsque vous appelez l'CreateDataSourceAPI.

- Type de base de données —Vous devez spécifier le type de base de données sous `sqlserver` la forme.
- Requête SQL : spécifiez des instructions de requête SQL telles que les opérations SELECT et JOIN. La taille des requêtes SQL doit être inférieure à 32 Ko. Amazon Kendra analysera tout le contenu de la base de données correspondant à votre requête.

  > **Note**
  >
  > Si le nom d'une table contient des caractères spéciaux (non alphanumériques), vous devez utiliser des crochets autour du nom de la table. Par exemple, *`select * from [my-database-table]`*

- Mode de synchronisation : spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation. Vous pouvez choisir entre :
  - `FORCED_FULL_CRAWL`pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
  - `FULL_CRAWL`pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
  - `CHANGE_LOG`pour indexer uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
- Nom de ressource Amazon (ARN) secret : indiquez le nom de ressource Amazon (ARN) d'un Secrets Manager secret contenant les informations d'authentification que vous avez créées dans votre Microsoft SQL Server compte. Le secret est stocké dans une structure JSON avec les clés suivantes :

```
{
    "user name": "database user name",
    "password": "password"
```

```
}
```

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

- IAM role —Spécifiez le `RoleArn` moment où vous appelez `CreateDataSource` pour fournir à un IAM rôle les autorisations d'accéder à votre Secrets Manager secret et d'appeler le public requis APIs pour le Microsoft SQL Server connecteur et Amazon Kendra. Pour plus d'informations, consultez la section IAM Rôles des sources de Microsoft SQL Server données.

Vous pouvez également ajouter les fonctionnalités optionnelles suivantes :

- Virtual Private Cloud (VPC) **`VpcConfiguration`** —Spécifiez le moment de votre appel. `CreateDataSource` Pour de plus amples informations, veuillez consulter Configuration Amazon Kendra pour utiliser un Amazon VPC.
- Filtres d'inclusion et d'exclusion : vous pouvez spécifier si vous souhaitez inclure un contenu spécifique en utilisant l'utilisateur IDs, les groupes, la source URLs, les horodatages et les fuseaux horaires.
- Filtrage du contexte utilisateur et contrôle d'accèsAmazon Kendra : analyse la liste de contrôle d'accès (ACL) de vos documents, si vous disposez d'une ACL pour vos documents. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.
- Mappages de champs : choisissez de mapper les champs de votre source de Microsoft SQL Server données à vos champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.

> **Note**
>
> Le champ du corps du document ou l'équivalent du corps du document pour vos documents est requis pour Amazon Kendra effectuer une recherche dans vos

documents. Vous devez associer le nom du champ du corps du document dans votre source de données au nom du champ d'index`_document_body`. Tous les autres champs sont facultatifs.

Pour obtenir la liste des autres clés JSON importantes à configurer, consultezSchéma de modèle Microsoft SQL Server.

Remarques

- Les lignes de base de données supprimées ne seront pas suivies lors de la Amazon Kendra vérification du contenu mis à jour.
- La taille des noms et des valeurs des champs d'une ligne de votre base de données ne peut pas dépasser 400 Ko.
- Si la source de données de votre base de données contient une grande quantité de données et que vous ne souhaitez pas Amazon Kendra indexer tout le contenu de votre base de données après la première synchronisation, vous pouvez choisir de ne synchroniser que les documents nouveaux, modifiés ou supprimés.
- Il est recommandé de fournir des informations d'identification Amazon Kendra de base de données en lecture seule.
- Il est recommandé d'éviter d'ajouter des tableaux contenant des données sensibles ou des informations personnelles identifiables (PII).

Microsoft Teams

Microsoft Teams est un outil de collaboration d'entreprise pour la messagerie, les réunions et le partage de fichiers. Si vous êtes un utilisateur de Microsoft Teams, vous pouvez l'utiliser Amazon Kendra pour indexer votre source de données Microsoft Teams.

Vous pouvez vous connecter Amazon Kendra à votre source de données Microsoft Teams à l'aide de la Amazon Kendra console et de l'TemplateConfigurationAPI.

Pour résoudre les problèmes liés à votre connecteur de source de données Amazon Kendra Microsoft Teams, consultezDépannage des sources de données.

Rubriques

- Fonctionnalités prises en charge
- Prérequis
- Instructions de connexion
- En savoir plus
- Remarques

Fonctionnalités prises en charge

- Mappages de champs
- Contrôle d'accès des utilisateurs
- Filtres d'inclusion/exclusion
- Synchronisation complète et incrémentielle du contenu
- Cloud privé virtuel (VPC)

Prérequis

Avant de pouvoir Amazon Kendra indexer votre source de données Microsoft Teams, apportez ces modifications à vos AWS comptes et à vos équipes Microsoft.

Dans Microsoft Teams, assurez-vous d'avoir :

- Création d'un compte Microsoft Teams dans Office 365.
- Vous avez noté votre identifiant de client Microsoft 365. Vous pouvez trouver votre ID de locataire dans les propriétés de votre portail Azure Active Directory ou dans votre OAuth application.
- J'ai configuré une OAuth application sur le portail Azure et noté l'ID du client et le secret du client ou les informations d'identification du client. Consultez le didacticiel Microsoft et l'exemple d'application enregistrée pour plus d'informations.

> **Note**
>
> Lorsque vous créez ou enregistrez une application sur le portail Azure, l'ID secret représente la valeur secrète réelle. Vous devez prendre note ou enregistrer la valeur secrète réelle immédiatement lors de la création du secret et de l'application. Vous pouvez accéder à votre secret en sélectionnant le nom de votre application sur le portail Azure, puis en accédant à l'option de menu sur les certificats et les secrets.

Vous pouvez accéder à votre ID client en sélectionnant le nom de votre application sur le portail Azure, puis en accédant à la page de présentation. L'ID de l'application (client) est l'ID du client.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

- Les autorisations nécessaires ont été ajoutées. Vous pouvez choisir d'ajouter toutes les autorisations ou de limiter la portée en sélectionnant moins d'autorisations en fonction des entités que vous souhaitez explorer. Le tableau suivant répertorie les autorisations au niveau de l'application par entité correspondante :

Entité	Autorisations requises pour la synchronisation des données	Autorisations requises pour Identity Sync
Publication sur la chaîne	• ChannelMessage. Lisez. Tout • Grouper. Tout lire • Utilisateur.Read • Utilisateur.Lisez.Tout	TeamMember. Lisez. Tout
Attachement au canal	• ChannelMessage. Lisez. Tout • Grouper. Tout lire • Utilisateur.Read • Utilisateur.Lisez.Tout	TeamMember. Lisez. Tout
Wiki de la chaîne	• Grouper. Tout lire • Utilisateur.Read	TeamMember. Lisez. Tout

Entité	Autorisations requises pour la synchronisation des données	Autorisations requises pour Identity Sync
	• Utilisateur.Lisez.Tout	
Message de chat	• Discuter. Tout lire • ChatMessage. Lisez. Tout • ChatMember. Lisez. Tout • Utilisateur.Read • Utilisateur.Lisez.Tout • Grouper. Tout lire	TeamMember. Lisez. Tout
Chat de réunion	• Discuter. Tout lire • ChatMessage.Lire • ChatMember. Lisez. Tout • Utilisateur.Read • Utilisateur.Lisez.Tout • Grouper. Tout lire	TeamMember. Lisez. Tout
Pièce jointe au chat	• Discuter. Tout lire • ChatMessage.Lire • ChatMember. Lisez. Tout • Utilisateur.Read • Utilisateur.Lisez.Tout • Grouper. Tout lire	TeamMember. Lisez. Tout
Dossier de réunion	• Discuter. Tout lire • ChatMessage. Lisez. Tout • ChatMember. Lisez. Tout • Utilisateur.Read • Utilisateur.Lisez.Tout • Grouper. Tout lire • Fichiers.Lisez.Tout	TeamMember. Lisez. Tout

Entité	Autorisations requises pour la synchronisation des données	Autorisations requises pour Identity Sync
Calendrier des réunions	• Discuter. Tout lire • ChatMessage. Lisez. Tout • ChatMember. Lisez. Tout • Utilisateur.Read • Utilisateur.Lisez.Tout • Grouper. Tout lire • Fichiers.Lisez.Tout	TeamMember. Lisez. Tout
Notes de réunion	• Utilisateur.Read • Utilisateur.Lisez.Tout • Grouper. Tout lire • Fichiers.Lisez.Tout	TeamMember. Lisez. Tout

- Il est vérifié que chaque document est unique dans Microsoft Teams et dans les autres sources de données que vous prévoyez d'utiliser pour le même index. Chaque source de données que vous souhaitez utiliser pour un index ne doit pas contenir le même document dans toutes les sources de données. IDs Les documents sont globaux par rapport à un index et doivent être uniques par index.

Dans votre Compte AWS, assurez-vous d'avoir :

- Création d'un Amazon Kendra index et, si vous utilisez l'API, notez l'ID de l'index.
- Vous avez créé un IAM rôle pour votre source de données et, si vous utilisez l'API, notez l'ARN du IAM rôle.

> **Note**
>
> Si vous modifiez votre type d'authentification et vos informations d'identification, vous devez mettre à jour votre IAM rôle pour accéder au bon identifiant AWS Secrets Manager secret.

- Stockez vos informations d'authentification Microsoft Teams dans un AWS Secrets Manager secret et, si vous utilisez l'API, notez l'ARN du secret.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

Si vous n'avez pas de IAM rôle ou de secret existant, vous pouvez utiliser la console pour créer un nouveau IAM rôle et un nouveau Secrets Manager secret lorsque vous connectez votre source de données Microsoft Teams à Amazon Kendra. Si vous utilisez l'API, vous devez fournir l'ARN d'un IAM rôle et d'un Secrets Manager secret existants, ainsi qu'un identifiant d'index.

Instructions de connexion

Pour vous connecter Amazon Kendra à votre source de données Microsoft Teams, vous devez fournir les informations nécessaires sur votre source de données Microsoft Teams afin de Amazon Kendra pouvoir accéder à vos données. Si vous n'avez pas encore configuré Microsoft Teams pour Amazon Kendra, consultezPrérequis.

Console

Pour vous connecter Amazon Kendra à Microsoft Teams

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la.
2. Dans le volet de navigation de gauche, choisissez Index, puis choisissez l'index que vous souhaitez utiliser dans la liste des index.

   > **Note**
   >
   > Vous pouvez choisir de configurer ou de modifier vos paramètres de contrôle d'accès utilisateur dans les paramètres de l'index.

3. Sur la page de démarrage, choisissez Ajouter une source de données.
4. Sur la page Ajouter une source de données, choisissez le connecteur Microsoft Teams, puis sélectionnez Ajouter un connecteur. Si vous utilisez la version 2 (le cas échéant), choisissez le connecteur Microsoft Teams avec la balise « V2.0 ».

5. Sur la page Spécifier les détails de la source de données, entrez les informations suivantes :

   a. Dans Nom et description, pour Nom de la source de données : entrez le nom de votre source de données. Vous pouvez inclure des tirets, mais pas des espaces.

   b. (Facultatif) Description : entrez une description facultative pour votre source de données.

   c. Dans la langue par défaut : choisissez une langue pour filtrer vos documents pour l'index. Sauf indication contraire, la langue par défaut est l'anglais. La langue spécifiée dans les métadonnées du document remplace la langue sélectionnée.

   d. Dans Balises, pour Ajouter une nouvelle balise : incluez des balises facultatives pour rechercher et filtrer vos ressources ou suivre vos AWS coûts.

   e. Choisissez Suivant.

6. Sur la page Définir l'accès et la sécurité, entrez les informations suivantes :

   a. ID de locataire —Entrez votre identifiant de locataire Microsoft 365. Vous pouvez trouver votre ID de locataire dans les propriétés de votre portail Azure Active Directory ou dans votre OAuth application.

   b. Autorisation : activez ou désactivez les informations de la liste de contrôle d'accès (ACL) pour vos documents, si vous disposez d'une ACL et que vous souhaitez l'utiliser pour le contrôle d'accès. L'ACL indique les documents auxquels les utilisateurs et les groupes peuvent accéder. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.

   c. AWS Secrets Manager secret —Choisissez un secret existant ou créez-en un nouveau Secrets Manager pour stocker vos informations d'authentification Microsoft Teams. Si vous choisissez de créer un nouveau secret, une fenêtre AWS Secrets Manager secrète s'ouvre.

      i. Entrez les informations suivantes dans la fenêtre Créer un AWS Secrets Manager secret :

         A. Nom secret : le nom de votre secret. Le préfixe « AmazonKendra -Microsoft Teams-» est automatiquement ajouté à votre nom secret.

         B. Pour l'ID client et le secret du client : entrez les informations d'authentification configurées dans Microsoft Teams sur le portail Azure.

      ii. Enregistrez et ajoutez votre secret.

d. Modèle de paiement : vous pouvez choisir un modèle de licence et de paiement pour votre compte Microsoft Teams. Les modèles de paiement du modèle A sont limités aux modèles de licence et de paiement qui nécessitent une conformité en matière de sécurité. Les modèles de paiement du modèle B conviennent aux modèles de licence et de paiement qui ne nécessitent pas de conformité en matière de sécurité.

e. Virtual Private Cloud (VPC) —Vous pouvez choisir d'utiliser un VPC. Dans ce cas, vous devez ajouter des sous-réseaux et des groupes de sécurité VPC.

f. Identity Crawler : spécifiez s'il faut activer l'explorateur Amazon Kendra d'identité. Le robot d'exploration d'identité utilise les informations de la liste de contrôle d'accès (ACL) de vos documents pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Si vous disposez d'une ACL pour vos documents et que vous choisissez de l'utiliser, vous pouvez également choisir d'activer le robot d'exploration Amazon Kendra d'identité pour configurer le filtrage contextuel utilisateur des résultats de recherche. Sinon, si le robot d'identification est désactivé, tous les documents peuvent être consultés publiquement. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'exploration d'identité est désactivé, vous pouvez également utiliser l'PutPrincipalMappingAPI pour télécharger les informations d'accès des utilisateurs et des groupes afin de filtrer le contexte utilisateur.

g. IAM rôle —Choisissez un IAM rôle existant ou créez-en un nouveau IAM pour accéder aux informations d'identification de votre référentiel et indexer le contenu.

> **Note**
>
> IAM les rôles utilisés pour les index ne peuvent pas être utilisés pour les sources de données. Si vous ne savez pas si un rôle existant est utilisé pour un index ou une FAQ, choisissez Créer un nouveau rôle pour éviter les erreurs.

h. Choisissez Suivant.

7. Sur la page Configurer les paramètres de synchronisation, entrez les informations suivantes :

a. Synchroniser le contenu : sélectionnez les types de contenu à analyser. Vous pouvez choisir d'explorer le contenu du chat, des équipes et du calendrier.

b. Configuration supplémentaire : spécifiez certaines dates de début et de fin du calendrier, les e-mails des utilisateurs, les noms des équipes, les noms des chaînes, les pièces jointes et OneNotes.

c. Mode de synchronisation : choisissez la manière dont vous souhaitez mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation.

   - Synchronisation complète : Indexez tout le contenu fraîchement, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
   - Nouvelle synchronisation modifiée : Indexez uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
   - Synchronisation nouvelle, modifiée ou supprimée : Indexez uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

d. Dans Calendrier d'exécution de la synchronisation, pour Fréquence : choisissez la fréquence à laquelle vous souhaitez synchroniser le contenu de votre source de données et mettre à jour votre index.

e. Choisissez Suivant.

8. Sur la page Définir les mappages de champs, entrez les informations suivantes :

   a. Champs de source de données par défaut : sélectionnez parmi les champs de source de données par défaut Amazon Kendra générés que vous souhaitez mapper à votre index.

   b. Ajouter un champ —Pour ajouter des champs de source de données personnalisés afin de créer un nom de champ d'index à mapper et le type de données du champ.

   c. Choisissez Suivant.

9. Sur la page Réviser et créer, vérifiez que les informations saisies sont correctes, puis sélectionnez Ajouter une source de données. Vous pouvez également choisir de modifier vos informations depuis cette page. Votre source de données apparaîtra sur la page Sources de données une fois que la source de données aura été ajoutée avec succès.

API

Pour vous connecter Amazon Kendra à Microsoft Teams

Vous devez spécifier un JSON du schéma de source de données à l'aide de l'TemplateConfigurationAPI. Vous devez fournir les informations suivantes :

- Source de données —Spécifiez le type de source de données tel que `MSTEAMS` lorsque vous utilisez le schéma TemplateConfigurationJSON. Spécifiez également la source de données `TEMPLATE` lorsque vous appelez l'CreateDataSourceAPI.
- ID de locataire : vous pouvez trouver votre identifiant de locataire dans les propriétés de votre portail Azure Active Directory ou dans votre OAuth application.
- Mode de synchronisation : spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation. Vous pouvez choisir entre :
  - `FORCED_FULL_CRAWL`pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
  - `FULL_CRAWL`pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
  - `CHANGE_LOG`pour indexer uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
- Nom de ressource Amazon (ARN) secret : indiquez le nom de ressource Amazon (ARN) d'un Secrets Manager secret contenant les informations d'authentification de votre compte Microsoft Teams. Le secret est stocké dans une structure JSON avec les clés suivantes :

```
{
    "clientId": "client ID",
    "clientSecret": "client secret"
}
```

- IAM role —Spécifiez à quel `RoleArn` moment vous appelez `CreateDataSource` pour fournir à un IAM rôle les autorisations d'accéder à votre Secrets Manager secret et pour appeler le public requis APIs pour le connecteur Microsoft Teams et Amazon Kendra. Pour plus d'informations, consultez la section IAM Rôles pour les sources de données Microsoft Teams.

Vous pouvez également ajouter les fonctionnalités optionnelles suivantes :

- Virtual Private Cloud (VPC) **`VpcConfiguration`** —Spécifiez le moment de votre appel. `CreateDataSource` Pour de plus amples informations, veuillez consulter Configuration Amazon Kendra pour utiliser un Amazon VPC.
- Types de documents/contenus : indiquez si vous souhaitez explorer les messages de chat et les pièces jointes, les publications et les pièces jointes des chaînes, les wikis des chaînes, le contenu du calendrier, les discussions de réunion, les fichiers et les notes.
- Contenu du calendrier : spécifiez une date de début et une heure de fin pour analyser le contenu du calendrier.
- Filtres d'inclusion et d'exclusion : spécifiez si vous souhaitez inclure ou exclure certains contenus dans Microsoft Teams. Vous pouvez inclure ou exclure les noms d'équipes, les noms de chaînes, les noms de fichiers et les types de fichiers, les e-mails des utilisateurs, OneNote les sections et OneNote les pages.

  > **Note**
  >
  > La plupart des sources de données utilisent des modèles d'expressions régulières, qui sont des modèles d'inclusion ou d'exclusion appelés filtres. Si vous spécifiez un filtre d'inclusion, seul le contenu correspondant au filtre d'inclusion est indexé. Tout document qui ne correspond pas au filtre d'inclusion n'est pas indexé. Si vous spécifiez un filtre d'inclusion et d'exclusion, les documents correspondant au filtre d'exclusion ne sont pas indexés, même s'ils correspondent au filtre d'inclusion.

- Identity Crawler : spécifiez s'il faut activer l'explorateur Amazon Kendra d'identité. Le robot d'exploration d'identité utilise les informations de la liste de contrôle d'accès (ACL) de vos documents pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Si vous disposez d'une ACL pour vos documents et que vous choisissez de l'utiliser, vous pouvez également choisir d'activer le robot d'exploration Amazon Kendra d'identité pour configurer le filtrage contextuel utilisateur des résultats de recherche. Sinon, si le robot d'identification est désactivé, tous les documents peuvent être consultés

publiquement. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'exploration d'identité est désactivé, vous pouvez également utiliser l'PutPrincipalMappingAPI pour télécharger les informations d'accès des utilisateurs et des groupes afin de filtrer le contexte utilisateur.

- Mappages de champs : choisissez de mapper les champs de votre source de données Microsoft Teams à vos champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.

> **Note**
>
> Le champ du corps du document ou l'équivalent du corps du document pour vos documents est requis pour Amazon Kendra effectuer une recherche dans vos documents. Vous devez associer le nom du champ du corps du document dans votre source de données au nom du champ d'index`_document_body`. Tous les autres champs sont facultatifs.

Pour obtenir la liste des autres clés JSON importantes à configurer, consultez le schéma du modèle Microsoft Teams.

En savoir plus

Pour en savoir plus sur l'intégration Amazon Kendra à votre source de données Microsoft Teams, consultez :

- Effectuez des recherches intelligentes dans la source de données Microsoft Teams de votre organisation à l'aide du Amazon Kendra connecteur pour Microsoft Teams

Remarques

- Lorsque les listes de contrôle d'accès (ACLs) sont activées, l'option « Synchroniser uniquement le contenu nouveau ou modifié » n'est pas disponible en raison des limites de l'API Microsoft Teams. Nous vous recommandons d'utiliser plutôt les modes « Synchronisation complète » ou « Synchronisation du contenu nouveau, modifié ou supprimé », ou de les désactiver ACLs si vous devez utiliser ce mode de synchronisation.

Microsoft Yammer

> **Note**
>
> Le connecteur Microsoft Yammer reste entièrement pris en charge pour les clients existants jusqu'au 31 mai 2026. Bien que ce connecteur ne soit plus disponible pour les nouveaux utilisateurs, les utilisateurs actuels peuvent continuer à l'utiliser sans interruption. Nous faisons constamment évoluer notre portefeuille de connecteurs afin de proposer des solutions plus évolutives et personnalisables. Pour les intégrations futures, nous vous recommandons d'explorer le framework de connecteurs personnalisés Amazon Kendra [1], conçu pour prendre en charge un plus large éventail de cas d'utilisation en entreprise avec une flexibilité accrue.

Microsoft Yammer est un outil de collaboration d'entreprise pour la messagerie, les réunions et le partage de fichiers. Si vous êtes un utilisateur de Microsoft Yammer, vous pouvez l'utiliser Amazon Kendra pour indexer votre source de données Microsoft Yammer.

Vous pouvez vous connecter Amazon Kendra à votre source de données Microsoft Yammer à l'aide de la Amazon Kendra console et de l'TemplateConfigurationAPI.

Pour résoudre les problèmes liés à votre connecteur de source de données Amazon Kendra Microsoft Yammer, consultezDépannage des sources de données.

Fonctionnalités prises en charge

- Mappages de champs
- Contrôle d'accès utilisateur
- Filtres d'inclusion/exclusion
- Synchronisation complète et incrémentielle du contenu
- Cloud privé virtuel (VPC)

Prérequis

Avant de pouvoir Amazon Kendra indexer votre source de données Microsoft Yammer, apportez ces modifications à votre compte et AWS à votre compte Microsoft Yammer.

Dans Microsoft Yammer, assurez-vous d'avoir :

- Création d'un compte administratif Microsoft Yammer dans Office 365.
- Vous avez noté votre nom d'utilisateur et votre mot de passe Microsoft Yammer.
- Vous avez noté votre identifiant de client Microsoft 365. Vous pouvez trouver votre ID de locataire dans les propriétés de votre portail Azure Active Directory ou dans votre OAuth application.
- J'ai configuré une OAuth application sur le portail Azure et noté l'ID du client et le secret du client ou les informations d'identification du client. Consultez le didacticiel Microsoft et l'exemple d'application enregistrée pour plus d'informations.

> **Note**
>
> Lorsque vous créez ou enregistrez une application sur le portail Azure, l'ID secret représente la valeur secrète réelle. Vous devez prendre note ou enregistrer la valeur secrète réelle immédiatement lors de la création du secret et de l'application. Vous pouvez accéder à votre secret en sélectionnant le nom de votre application sur le portail Azure, puis en accédant à l'option de menu sur les certificats et les secrets.
> Vous pouvez accéder à votre ID client en sélectionnant le nom de votre application sur le portail Azure, puis en accédant à la page de présentation. L'ID de l'application (client) est l'ID du client.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

- Il est vérifié que chaque document est unique dans Microsoft Yammer et dans les autres sources de données que vous prévoyez d'utiliser pour le même index. Chaque source de données que vous souhaitez utiliser pour un index ne doit pas contenir le même document dans toutes les sources de données. IDs Les documents sont globaux par rapport à un index et doivent être uniques par index.

Dans votre Compte AWS, assurez-vous d'avoir :

- Création d'un Amazon Kendra index et, si vous utilisez l'API, notez l'ID de l'index.
- Vous avez créé un IAM rôle pour votre source de données et, si vous utilisez l'API, notez l'ARN du IAM rôle.

> **Note**
>
> Si vous modifiez votre type d'authentification et vos informations d'identification, vous devez mettre à jour votre IAM rôle pour accéder au bon identifiant AWS Secrets Manager secret.

- Stockez vos informations d'authentification Microsoft Yammer dans un AWS Secrets Manager secret et, si vous utilisez l'API, notez l'ARN du secret.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

Si vous n'avez pas de IAM rôle ou de secret existant, vous pouvez utiliser la console pour créer un nouveau IAM rôle et un nouveau Secrets Manager secret lorsque vous connectez votre source de données Microsoft Yammer à Amazon Kendra. Si vous utilisez l'API, vous devez fournir l'ARN d'un IAM rôle et d'un Secrets Manager secret existants, ainsi qu'un identifiant d'index.

Instructions de connexion

Pour vous connecter Amazon Kendra à votre source de données Microsoft Yammer, vous devez fournir les informations nécessaires sur votre source de données Microsoft Yammer afin de Amazon Kendra pouvoir accéder à vos données. Si vous n'avez pas encore configuré Microsoft Yammer pour Amazon Kendra, consultezPrérequis.

Console

Pour vous connecter Amazon Kendra à Microsoft Yammer

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la.

2. Dans le volet de navigation de gauche, choisissez Index, puis choisissez l'index que vous souhaitez utiliser dans la liste des index.

> **Note**
>
> Vous pouvez choisir de configurer ou de modifier vos paramètres de contrôle d'accès utilisateur dans les paramètres de l'index.

3. Sur la page de démarrage, choisissez Ajouter une source de données.
4. Sur la page Ajouter une source de données, choisissez le connecteur Microsoft Yammer, puis sélectionnez Ajouter un connecteur. Si vous utilisez la version 2 (le cas échéant), choisissez le connecteur Microsoft Yammer avec la balise « V2.0 ».
5. Sur la page Spécifier les détails de la source de données, entrez les informations suivantes :

   a. Dans Nom et description, pour Nom de la source de données : entrez le nom de votre source de données. Vous pouvez inclure des tirets, mais pas des espaces.

   b. (Facultatif) Description : entrez une description facultative pour votre source de données.

   c. Dans la langue par défaut : choisissez une langue pour filtrer vos documents pour l'index. Sauf indication contraire, la langue par défaut est l'anglais. La langue spécifiée dans les métadonnées du document remplace la langue sélectionnée.

   d. Dans Balises, pour Ajouter une nouvelle balise : incluez des balises facultatives pour rechercher et filtrer vos ressources ou suivre vos AWS coûts.

   e. Choisissez Suivant.

6. Sur la page Définir l'accès et la sécurité, entrez les informations suivantes :

   a. Autorisation : activez ou désactivez les informations de la liste de contrôle d'accès (ACL) pour vos documents, si vous disposez d'une ACL et que vous souhaitez l'utiliser pour le contrôle d'accès. L'ACL indique les documents auxquels les utilisateurs et les groupes peuvent accéder. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.

   b. AWS Secrets Manager secret —Choisissez un secret existant ou créez-en un nouveau Secrets Manager pour stocker vos informations d'authentification Microsoft Yammer. Si vous choisissez de créer un nouveau secret, une fenêtre AWS Secrets Manager secrète s'ouvre.

i. Entrez les informations suivantes dans la fenêtre Créer un AWS Secrets Manager secret :

A. Nom secret : le nom de votre secret. Le préfixe « AmazonKendra -Microsoft Yammer » est automatiquement ajouté à votre nom secret.

B. Pour Nom d'utilisateur, mot de passe : entrez votre nom d'utilisateur et votre mot de passe Microsoft Yammer.

C. Pour l'ID client, secret du client : entrez les informations d'authentification configurées dans Microsoft Yammer sur le portail Azure.

ii. Enregistrez et ajoutez votre secret.

c. Virtual Private Cloud (VPC) —Vous pouvez choisir d'utiliser un VPC. Dans ce cas, vous devez ajouter des sous-réseaux et des groupes de sécurité VPC.

d. Identity Crawler : spécifiez s'il faut activer l'explorateur Amazon Kendra d'identité. Le robot d'exploration d'identité utilise les informations de la liste de contrôle d'accès (ACL) de vos documents pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Si vous disposez d'une ACL pour vos documents et que vous choisissez de l'utiliser, vous pouvez également choisir d'activer le robot d'exploration Amazon Kendra d'identité pour configurer le filtrage contextuel utilisateur des résultats de recherche. Sinon, si le robot d'identification est désactivé, tous les documents peuvent être consultés publiquement. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'exploration d'identité est désactivé, vous pouvez également utiliser l'PutPrincipalMappingAPI pour télécharger les informations d'accès des utilisateurs et des groupes afin de filtrer le contexte utilisateur.

e. IAM rôle —Choisissez un IAM rôle existant ou créez-en un nouveau IAM pour accéder aux informations d'identification de votre référentiel et indexer le contenu.

> **Note**
>
> IAM les rôles utilisés pour les index ne peuvent pas être utilisés pour les sources de données. Si vous ne savez pas si un rôle existant est utilisé pour un index ou une FAQ, choisissez Créer un nouveau rôle pour éviter les erreurs.

f. Choisissez Suivant.

7. Sur la page Configurer les paramètres de synchronisation, entrez les informations suivantes :

a. Depuis la date : spécifiez la date de début de l'analyse de vos données dans Microsoft Yammer.

b. Synchroniser le contenu : sélectionnez le type de contenu à explorer. Par exemple, les messages publics, les messages privés et les pièces jointes.

c. Configuration supplémentaire : spécifiez certains noms de communauté que vous souhaitez explorer et utilisez également des modèles d'expressions régulières pour inclure ou exclure certains contenus.

d. Mode de synchronisation : choisissez la manière dont vous souhaitez mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation.

   - Synchronisation complète : Indexez tout le contenu fraîchement, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
   - Nouvelle synchronisation modifiée : Indexez uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
   - Synchronisation nouvelle, modifiée ou supprimée : Indexez uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

e. Dans Calendrier d'exécution de la synchronisation, pour Fréquence : choisissez la fréquence à laquelle vous souhaitez synchroniser le contenu de votre source de données et mettre à jour votre index.

f. Choisissez Suivant.

8. Sur la page Définir les mappages de champs, entrez les informations suivantes :

   a. Champs de source de données par défaut : sélectionnez parmi les champs de source de données par défaut Amazon Kendra générés que vous souhaitez mapper à votre index.

b. Ajouter un champ —Pour ajouter des champs de source de données personnalisés afin de créer un nom de champ d'index à mapper et le type de données du champ.

c. Choisissez Suivant.

9. Sur la page Réviser et créer, vérifiez que les informations saisies sont correctes, puis sélectionnez Ajouter une source de données. Vous pouvez également choisir de modifier vos informations depuis cette page. Votre source de données apparaîtra sur la page Sources de données une fois que la source de données aura été ajoutée avec succès.

API

Pour vous connecter Amazon Kendra à Microsoft Yammer

Vous devez spécifier un JSON du schéma de source de données à l'aide de l'TemplateConfigurationAPI. Vous devez fournir les informations suivantes :

- Source de données —Spécifiez le type de source de données tel que `YAMMER` lorsque vous utilisez le schéma TemplateConfigurationJSON. Spécifiez également la source de données `TEMPLATE` lorsque vous appelez l'CreateDataSourceAPI.
- Mode de synchronisation : spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation. Vous pouvez choisir entre :
  - `FORCED_FULL_CRAWL`pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
  - `FULL_CRAWL`pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
  - `CHANGE_LOG`pour indexer uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

- Nom de ressource Amazon (ARN) secret : indiquez le nom de ressource Amazon (ARN) d'un Secrets Manager secret contenant les informations d'authentification de votre compte Microsoft Yammer. Le secret est stocké dans une structure JSON avec les clés suivantes :

```
{
    "username": "user name",
    "password": "password",
    "clientId": "client ID",
    "clientSecret": "client secret"
}
```

- IAM role —Spécifiez à quel `RoleArn` moment vous appelez `CreateDataSource` pour fournir à un IAM rôle les autorisations d'accéder à votre Secrets Manager secret et pour appeler le public requis APIs pour le connecteur Microsoft Yammer et. Amazon Kendra Pour plus d'informations, consultez la section IAM Rôles des sources de données Microsoft Yammer.

Vous pouvez également ajouter les fonctionnalités optionnelles suivantes :

- Virtual Private Cloud (VPC) **`VpcConfiguration`** —Spécifiez le moment de votre appel. `CreateDataSource` Pour de plus amples informations, veuillez consulter Configuration Amazon Kendra pour utiliser un Amazon VPC.
- Types de documents/contenus : indiquez si vous souhaitez analyser le contenu de la communauté, les messages et pièces jointes, ainsi que les messages privés.
- Filtres d'inclusion et d'exclusion : spécifiez s'il faut inclure ou exclure certains contenus.

> **Note**
>
> La plupart des sources de données utilisent des modèles d'expressions régulières, qui sont des modèles d'inclusion ou d'exclusion appelés filtres. Si vous spécifiez un filtre d'inclusion, seul le contenu correspondant au filtre d'inclusion est indexé. Tout document qui ne correspond pas au filtre d'inclusion n'est pas indexé. Si vous spécifiez un filtre d'inclusion et d'exclusion, les documents correspondant au filtre d'exclusion ne sont pas indexés, même s'ils correspondent au filtre d'inclusion.

- Identity Crawler : spécifiez s'il faut activer l'explorateur Amazon Kendra d'identité. Le robot d'exploration d'identité utilise les informations de la liste de contrôle d'accès (ACL) de vos documents pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Si vous disposez d'une ACL pour vos documents et que vous

choisissez de l'utiliser, vous pouvez également choisir d'activer le robot d'exploration Amazon Kendra d'identité pour configurer le filtrage contextuel utilisateur des résultats de recherche. Sinon, si le robot d'identification est désactivé, tous les documents peuvent être consultés publiquement. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'exploration d'identité est désactivé, vous pouvez également utiliser l'PutPrincipalMappingAPI pour télécharger les informations d'accès des utilisateurs et des groupes afin de filtrer le contexte utilisateur.

- Mappages de champs : choisissez de mapper les champs de votre source de données Microsoft Yammer à vos Amazon Kendra champs d'index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.

> **Note**
>
> Le champ du corps du document ou l'équivalent du corps du document pour vos documents est requis pour Amazon Kendra effectuer une recherche dans vos documents. Vous devez associer le nom du champ du corps du document dans votre source de données au nom du champ d'index`_document_body`. Tous les autres champs sont facultatifs.

Pour obtenir la liste des autres clés JSON importantes à configurer, consultez le schéma du modèle Microsoft Yammer.

En savoir plus

Pour en savoir plus sur l'intégration Amazon Kendra à votre source de données Microsoft Yammer, consultez :

- Annonce du connecteur Yammer pour Amazon Kendra

MySQL

> **Note**
>
> MySQLLe connecteur reste entièrement pris en charge pour les clients existants jusqu'au 31 mai 2026. Bien que ce connecteur ne soit plus disponible pour les nouveaux utilisateurs, les utilisateurs actuels peuvent continuer à l'utiliser sans interruption. Nous faisons constamment

évoluer notre portefeuille de connecteurs afin de proposer des solutions plus évolutives et personnalisables. Pour les intégrations futures, nous vous recommandons d'explorer le framework de connecteurs personnalisés Amazon Kendra [1], conçu pour prendre en charge un plus large éventail de cas d'utilisation en entreprise avec une flexibilité accrue.

MySQLest un système de gestion de base de données relationnelle open source. Si vous êtes un MySQL utilisateur, vous pouvez l'utiliser Amazon Kendra pour indexer votre source de MySQL données. Le connecteur Amazon Kendra MySQL de source de données prend en charge MySQL 8.0. 21.

Vous pouvez vous connecter Amazon Kendra à votre source de MySQL données à l'aide de la Amazon Kendra console et de l'TemplateConfigurationAPI.

Pour résoudre les problèmes liés à votre connecteur de source de Amazon Kendra MySQL données, consultezDépannage des sources de données.

Rubriques

- Fonctionnalités prises en charge
- Prérequis
- Instructions de connexion
- Remarques

Fonctionnalités prises en charge

- Mappages de champs
- Filtrage du contexte utilisateur
- Filtres d'inclusion/exclusion
- Synchronisation complète et incrémentielle du contenu
- Cloud privé virtuel (VPC)

Prérequis

Avant de pouvoir Amazon Kendra indexer votre source de MySQL données, apportez ces modifications à vos AWS comptes MySQL and.

DansMySQL, assurez-vous d'avoir :

- Notez le nom d'utilisateur et le mot de passe de votre base de données.

> **Important**
>
> Il est recommandé de fournir des informations d'identification Amazon Kendra de base de données en lecture seule.

- Vous avez copié l'URL, le port et l'instance de votre hôte de base de données.
- Il est vérifié que chaque document est unique dans MySQL et entre les autres sources de données que vous prévoyez d'utiliser pour le même index. Chaque source de données que vous souhaitez utiliser pour un index ne doit pas contenir le même document dans toutes les sources de données. IDs Les documents sont globaux par rapport à un index et doivent être uniques par index.

Dans votre Compte AWS, assurez-vous d'avoir :

- Création d'un Amazon Kendra index et, si vous utilisez l'API, notez l'ID de l'index.
- Vous avez créé un IAM rôle pour votre source de données et, si vous utilisez l'API, notez l'ARN du IAM rôle.

> **Note**
>
> Si vous modifiez votre type d'authentification et vos informations d'identification, vous devez mettre à jour votre IAM rôle pour accéder au bon identifiant AWS Secrets Manager secret.

- Stockez vos informations d'MySQLauthentification dans un AWS Secrets Manager secret et, si vous utilisez l'API, notez l'ARN du secret.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

Si vous n'avez pas de IAM rôle ou de secret existant, vous pouvez utiliser la console pour créer un nouveau IAM rôle et un nouveau Secrets Manager secret lorsque vous connectez votre source de MySQL données à Amazon Kendra. Si vous utilisez l'API, vous devez fournir l'ARN d'un IAM rôle et d'un Secrets Manager secret existants, ainsi qu'un identifiant d'index.

Instructions de connexion

Pour vous connecter Amazon Kendra à votre source de MySQL données, vous devez fournir les détails de vos MySQL informations d'identification afin de Amazon Kendra pouvoir accéder à vos données. Si vous n'avez pas encore configuré MySQL pour Amazon Kendra voirPrérequis.

Console

Pour vous connecter Amazon Kendra à MySQL

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la.
2. Dans le volet de navigation de gauche, choisissez Index, puis choisissez l'index que vous souhaitez utiliser dans la liste des index.

> **Note**
>
> Vous pouvez choisir de configurer ou de modifier vos paramètres de contrôle d'accès utilisateur dans les paramètres de l'index.

3. Sur la page de démarrage, choisissez Ajouter une source de données.
4. Sur la page Ajouter une source de données, choisissez MySQLconnecteur, puis sélectionnez Ajouter un connecteur. Si vous utilisez la version 2 (le cas échéant), choisissez le MySQLconnecteur avec la balise « V2.0 ».
5. Sur la page Spécifier les détails de la source de données, entrez les informations suivantes :

   a. Dans Nom et description, pour Nom de la source de données : entrez le nom de votre source de données. Vous pouvez inclure des tirets, mais pas des espaces.
   b. (Facultatif) Description : entrez une description facultative pour votre source de données.
   c. Dans la langue par défaut : choisissez une langue pour filtrer vos documents pour l'index. Sauf indication contraire, la langue par défaut est l'anglais. La langue spécifiée dans les métadonnées du document remplace la langue sélectionnée.
   d. Dans Balises, pour Ajouter une nouvelle balise : incluez des balises facultatives pour rechercher et filtrer vos ressources ou suivre vos AWS coûts.

e. Choisissez Suivant.

6. Sur la page Définir l'accès et la sécurité, entrez les informations suivantes :

   a. Dans Source, entrez les informations suivantes :

   b. Hôte : entrez le nom d'hôte de la base de données.

   c. Port — Entrez le port de base de données.

   d. Instance — Entrez l'instance de base de données.

   e. Activer l'emplacement du certificat SSL : choisissez d'entrer le Amazon S3 chemin d'accès à votre fichier de certificat SSL.

   f. Dans Authentification, entrez les informations suivantes :

      - AWS Secrets Manager secret —Choisissez un secret existant ou créez-en un nouveau Secrets Manager pour stocker vos informations MySQL d'authentification. Si vous choisissez de créer un nouveau secret, une fenêtre AWS Secrets Manager secrète s'ouvre.

        A. Entrez les informations suivantes dans la fenêtre Créer un AWS Secrets Manager secret :

           I. Nom secret : le nom de votre secret. Le préfixe « AmazonKendra - MySQL - » est automatiquement ajouté à votre nom secret.

           II. Pour le nom d'utilisateur et le mot de passe de la base de données, entrez les valeurs d'identification d'authentification que vous avez copiées depuis votre base de données.

        B. Choisissez Enregistrer.

   g. Virtual Private Cloud (VPC) —Vous pouvez choisir d'utiliser un VPC. Dans ce cas, vous devez ajouter des sous-réseaux et des groupes de sécurité VPC.

   h. IAM rôle —Choisissez un IAM rôle existant ou créez-en un nouveau IAM pour accéder aux informations d'identification de votre référentiel et indexer le contenu.

> **Note**
>
> IAM les rôles utilisés pour les index ne peuvent pas être utilisés pour les sources de données. Si vous ne savez pas si un rôle existant est utilisé pour un index ou une FAQ, choisissez Créer un nouveau rôle pour éviter les erreurs.

i. Choisissez Suivant.

7. Sur la page Configurer les paramètres de synchronisation, entrez les informations suivantes :

   a. Dans le champ d'application de la synchronisation, choisissez l'une des options suivantes :

      - Requête SQL —Entrez des instructions de requête SQL telles que les opérations SELECT et JOIN. La taille des requêtes SQL doit être inférieure à 32 Ko. Amazon Kendra analysera tout le contenu de la base de données correspondant à votre requête.
      - Colonne clé primaire : indiquez la clé primaire pour la table de base de données. Cela identifie une table au sein de votre base de données.
      - Colonne de titre : indiquez le nom de la colonne de titre du document dans votre table de base de données.
      - Colonne du corps : indiquez le nom de la colonne du corps du document dans votre table de base de données.

   b. Dans Configuration supplémentaire (facultatif), choisissez l'une des options suivantes pour synchroniser un contenu spécifique au lieu de synchroniser tous les fichiers :

      - Colonnes détectant les modifications : entrez le nom des colonnes qui Amazon Kendra seront utilisées pour détecter les modifications de contenu. Amazon Kendra réindexera le contenu en cas de modification de l'une de ces colonnes.
      - IDsColonne des utilisateurs : entrez le nom de la colonne contenant l'utilisateur devant être autorisé IDs à accéder au contenu.
      - Colonne Groupes : entrez le nom de la colonne contenant les groupes autorisés à accéder au contenu.
      - URLsColonne source —Entrez le nom de la colonne contenant la source URLs à indexer.
      - Colonne horodatage : entrez le nom de la colonne contenant les horodatages. Amazon Kendra utilise les informations d'horodatage pour détecter les modifications apportées à votre contenu et synchroniser uniquement le contenu modifié.
      - Colonne fuseaux horaires : entrez le nom de la colonne contenant les fuseaux horaires du contenu à analyser.
      - Format d'horodatage : entrez le nom de la colonne contenant les formats d'horodatage à utiliser pour détecter les modifications de contenu et resynchroniser votre contenu.

c. Mode de synchronisation : choisissez la manière dont vous souhaitez mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation.

   - Synchronisation complète : Indexez tout le contenu fraîchement, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
   - Nouvelle synchronisation modifiée : indexez uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
   - Synchronisation nouvelle, modifiée ou supprimée : Indexez uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

d. Dans Synchroniser le calendrier d'exécution, pour Fréquence : à quelle fréquence Amazon Kendra sera synchronisée avec votre source de données.

e. Choisissez Suivant.

8. Sur la page Définir les mappages de champs, entrez les informations suivantes :

   a. Sélectionnez parmi les champs de source de données par défaut générés (Document IDs, Titres des documents et Source URLs) que vous souhaitez Amazon Kendra mapper à l'index.

   b. Ajouter un champ —Pour ajouter des champs de source de données personnalisés afin de créer un nom de champ d'index à mapper et le type de données du champ.

   c. Choisissez Suivant.

9. Sur la page Réviser et créer, vérifiez que les informations saisies sont correctes, puis sélectionnez Ajouter une source de données. Vous pouvez également choisir de modifier vos informations depuis cette page. Votre source de données apparaîtra sur la page Sources de données une fois qu'elle aura été ajoutée avec succès.

API

Pour vous connecter Amazon Kendra à MySQL

Vous devez spécifier les éléments suivants à l'aide de l'TemplateConfigurationAPI :

- Source de données —Spécifiez le type de source de données tel que `JDBC` lorsque vous utilisez le schéma TemplateConfigurationJSON. Spécifiez également la source de données `TEMPLATE` lorsque vous appelez l'CreateDataSourceAPI.
- Type de base de données —Vous devez spécifier le type de base de données sous `mySql` la forme.
- Requête SQL : spécifiez des instructions de requête SQL telles que les opérations SELECT et JOIN. La taille des requêtes SQL doit être inférieure à 32 Ko. Amazon Kendra analysera tout le contenu de la base de données correspondant à votre requête.
- Mode de synchronisation : spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation. Vous pouvez choisir entre :
  - `FORCED_FULL_CRAWL`pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
  - `FULL_CRAWL`pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
  - `CHANGE_LOG`pour indexer uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
- Nom de ressource Amazon (ARN) secret : indiquez le nom de ressource Amazon (ARN) d'un Secrets Manager secret contenant les informations d'authentification que vous avez créées dans votre MySQL compte. Le secret est stocké dans une structure JSON avec les clés suivantes :

```
{
    "user name": "database user name",
```

```
    "password": "password"
}
```

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

- IAM role —Spécifiez le `RoleArn` moment où vous appelez `CreateDataSource` pour fournir à un IAM rôle les autorisations d'accéder à votre Secrets Manager secret et d'appeler le public requis APIs pour le MySQL connecteur et Amazon Kendra. Pour plus d'informations, consultez la section IAM Rôles des sources de MySQL données.

Vous pouvez également ajouter les fonctionnalités optionnelles suivantes :

- Virtual Private Cloud (VPC) **`VpcConfiguration`** —Spécifiez le moment de votre appel. `CreateDataSource` Pour de plus amples informations, veuillez consulter Configuration Amazon Kendra pour utiliser un Amazon VPC.
- Filtres d'inclusion et d'exclusion : vous pouvez spécifier si vous souhaitez inclure un contenu spécifique en utilisant l'utilisateur IDs, les groupes, la source URLs, les horodatages et les fuseaux horaires.
- Filtrage du contexte utilisateur et contrôle d'accèsAmazon Kendra : analyse la liste de contrôle d'accès (ACL) de vos documents, si vous disposez d'une ACL pour vos documents. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.
- Mappages de champs : choisissez de mapper les champs de votre source de MySQL données à vos champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.

> **Note**
>
> Le champ du corps du document ou l'équivalent du corps du document pour vos documents est requis pour Amazon Kendra effectuer une recherche dans vos

documents. Vous devez associer le nom du champ du corps du document dans votre source de données au nom du champ d'index`_document_body`. Tous les autres champs sont facultatifs.

Remarques

- Les lignes de base de données supprimées ne seront pas suivies lors de la Amazon Kendra vérification du contenu mis à jour.
- La taille des noms et des valeurs des champs d'une ligne de votre base de données ne peut pas dépasser 400 Ko.
- Si la source de données de votre base de données contient une grande quantité de données et que vous ne souhaitez pas Amazon Kendra indexer tout le contenu de votre base de données après la première synchronisation, vous pouvez choisir de ne synchroniser que les documents nouveaux, modifiés ou supprimés.
- Il est recommandé de fournir des informations d'identification Amazon Kendra de base de données en lecture seule.
- Il est recommandé d'éviter d'ajouter des tableaux contenant des données sensibles ou des informations personnelles identifiables (PII).

Oracle Database

> **Note**
>
> Oracle DatabaseLe connecteur reste entièrement pris en charge pour les clients existants jusqu'au 31 mai 2026. Bien que ce connecteur ne soit plus disponible pour les nouveaux utilisateurs, les utilisateurs actuels peuvent continuer à l'utiliser sans interruption. Nous faisons constamment évoluer notre portefeuille de connecteurs afin de proposer des solutions plus évolutives et personnalisables. Pour les intégrations futures, nous vous recommandons d'explorer le framework de connecteurs personnalisés Amazon Kendra [1], conçu pour prendre en charge un plus large éventail de cas d'utilisation en entreprise avec une flexibilité accrue.

Oracle Databaseest un système de gestion de base de données. Si vous êtes un Oracle Database utilisateur, vous pouvez l'utiliser Amazon Kendra pour indexer votre source de Oracle Database

données. Le connecteur de source de Amazon Kendra Oracle Database données prend en charge les bases de données Oracle 18c, 19c et 21c.

Vous pouvez vous connecter Amazon Kendra à votre source de Oracle Database données à l'aide de la Amazon Kendra console et de l'TemplateConfigurationAPI.

Pour résoudre les problèmes liés à votre connecteur de source de Amazon Kendra Oracle Database données, consultezDépannage des sources de données.

Rubriques

- Fonctionnalités prises en charge
- Prérequis
- Instructions de connexion
- Remarques

Fonctionnalités prises en charge

- Mappages de champs
- Filtrage du contexte utilisateur
- Filtres d'inclusion/exclusion
- Synchronisation complète et incrémentielle du contenu
- Cloud privé virtuel (VPC)

Prérequis

Avant de pouvoir Amazon Kendra indexer votre source de Oracle Database données, apportez ces modifications à vos AWS comptes Oracle Database and.

DansOracle Database, assurez-vous d'avoir :

- Notez le nom d'utilisateur et le mot de passe de votre base de données.

> **Important**
>
> Il est recommandé de fournir des informations d'identification Amazon Kendra de base de données en lecture seule.

- Vous avez copié l'URL, le port et l'instance de votre hôte de base de données.
- Il est vérifié que chaque document est unique dans Oracle Database et entre les autres sources de données que vous prévoyez d'utiliser pour le même index. Chaque source de données que vous souhaitez utiliser pour un index ne doit pas contenir le même document dans toutes les sources de données. IDs Les documents sont globaux par rapport à un index et doivent être uniques par index.

Dans votre Compte AWS, assurez-vous d'avoir :

- Création d'un Amazon Kendra index et, si vous utilisez l'API, notez l'ID de l'index.
- Vous avez créé un IAM rôle pour votre source de données et, si vous utilisez l'API, notez l'ARN du IAM rôle.

> **Note**
>
> Si vous modifiez votre type d'authentification et vos informations d'identification, vous devez mettre à jour votre IAM rôle pour accéder au bon identifiant AWS Secrets Manager secret.

- Stockez vos informations d'Oracle Databaseauthentification dans un AWS Secrets Manager secret et, si vous utilisez l'API, notez l'ARN du secret.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

Si vous n'avez pas de IAM rôle ou de secret existant, vous pouvez utiliser la console pour créer un nouveau IAM rôle et un nouveau Secrets Manager secret lorsque vous connectez votre source de Oracle Database données à Amazon Kendra. Si vous utilisez l'API, vous devez fournir l'ARN d'un IAM rôle et d'un Secrets Manager secret existants, ainsi qu'un identifiant d'index.

Instructions de connexion

Pour vous connecter Amazon Kendra à votre source de Oracle Database données, vous devez fournir les détails de vos Oracle Database informations d'identification afin de Amazon Kendra pouvoir accéder à vos données. Si vous n'avez pas encore configuré Oracle Database pour Amazon Kendra voirPrérequis.

Console

Pour vous connecter Amazon Kendra à Oracle Database

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la.
2. Dans le volet de navigation de gauche, choisissez Index, puis choisissez l'index que vous souhaitez utiliser dans la liste des index.

> **Note**
>
> Vous pouvez choisir de configurer ou de modifier vos paramètres de contrôle d'accès utilisateur dans les paramètres de l'index.

3. Sur la page de démarrage, choisissez Ajouter une source de données.
4. Sur la page Ajouter une source de données, choisissez Oracle Databaseconnecteur, puis sélectionnez Ajouter un connecteur. Si vous utilisez la version 2 (le cas échéant), choisissez le Oracle Databaseconnecteur avec la balise « V2.0 ».
5. Sur la page Spécifier les détails de la source de données, entrez les informations suivantes :
   a. Dans Nom et description, pour Nom de la source de données : entrez le nom de votre source de données. Vous pouvez inclure des tirets, mais pas des espaces.
   b. (Facultatif) Description : entrez une description facultative pour votre source de données.
   c. Dans la langue par défaut : choisissez une langue pour filtrer vos documents pour l'index. Sauf indication contraire, la langue par défaut est l'anglais. La langue spécifiée dans les métadonnées du document remplace la langue sélectionnée.
   d. Dans Balises, pour Ajouter une nouvelle balise : incluez des balises facultatives pour rechercher et filtrer vos ressources ou suivre vos AWS coûts.
   e. Choisissez Suivant.
6. Sur la page Définir l'accès et la sécurité, entrez les informations suivantes :

a. Dans Source, entrez les informations suivantes :

b. Hôte : entrez le nom d'hôte de la base de données.

c. Port — Entrez le port de base de données.

d. Instance — Entrez l'instance de base de données.

e. Activer l'emplacement du certificat SSL : choisissez d'entrer le Amazon S3 chemin d'accès à votre fichier de certificat SSL.

f. Dans Authentification, entrez les informations suivantes :

- AWS Secrets Manager secret —Choisissez un secret existant ou créez-en un nouveau Secrets Manager pour stocker vos informations Oracle Database d'authentification. Si vous choisissez de créer un nouveau secret, une fenêtre AWS Secrets Manager secrète s'ouvre.

  A. Entrez les informations suivantes dans la fenêtre Créer un AWS Secrets Manager secret :

     I. Nom secret : le nom de votre secret. Le préfixe « AmazonKendra - Oracle Database - » est automatiquement ajouté à votre nom secret.

     II. Pour le nom d'utilisateur et le mot de passe de la base de données, entrez les valeurs d'identification d'authentification que vous avez copiées depuis votre base de données.

  B. Choisissez Enregistrer.

g. Virtual Private Cloud (VPC) —Vous pouvez choisir d'utiliser un VPC. Dans ce cas, vous devez ajouter des sous-réseaux et des groupes de sécurité VPC.

h. IAM rôle —Choisissez un IAM rôle existant ou créez-en un nouveau IAM pour accéder aux informations d'identification de votre référentiel et indexer le contenu.

> **Note**
>
> IAM les rôles utilisés pour les index ne peuvent pas être utilisés pour les sources de données. Si vous ne savez pas si un rôle existant est utilisé pour un index ou une FAQ, choisissez Créer un nouveau rôle pour éviter les erreurs.

i. Choisissez Suivant.

7. Sur la page Configurer les paramètres de synchronisation, entrez les informations suivantes :

a. Dans le champ d'application de la synchronisation, choisissez l'une des options suivantes :

   - Requête SQL —Entrez des instructions de requête SQL telles que les opérations SELECT et JOIN. La taille des requêtes SQL doit être inférieure à 32 Ko. Amazon Kendra analysera tout le contenu de la base de données correspondant à votre requête.
   - Colonne clé primaire : indiquez la clé primaire pour la table de base de données. Cela identifie une table dans votre base de données.
   - Colonne de titre : indiquez le nom de la colonne de titre du document dans votre table de base de données.
   - Colonne du corps : indiquez le nom de la colonne du corps du document dans votre table de base de données.

b. Dans Configuration supplémentaire — facultatif, choisissez l'une des options suivantes pour synchroniser un contenu spécifique au lieu de synchroniser tous les fichiers :

   - Colonnes détectant les modifications : entrez le nom des colonnes qui Amazon Kendra seront utilisées pour détecter les modifications de contenu. Amazon Kendra réindexera le contenu en cas de modification de l'une de ces colonnes.
   - IDsColonne utilisateur : entrez le nom de la colonne contenant l'utilisateur devant être autorisé IDs à accéder au contenu.
   - Colonne Groupes : entrez le nom de la colonne contenant les groupes autorisés à accéder au contenu.
   - URLsColonne source —Entrez le nom de la colonne contenant la source URLs à indexer.
   - Colonne horodatage : entrez le nom de la colonne contenant les horodatages. Amazon Kendra utilise les informations d'horodatage pour détecter les modifications apportées à votre contenu et synchroniser uniquement le contenu modifié.
   - Colonne fuseaux horaires : entrez le nom de la colonne contenant les fuseaux horaires du contenu à analyser.
   - Format d'horodatage : entrez le nom de la colonne contenant les formats d'horodatage à utiliser pour détecter les modifications de contenu et resynchroniser votre contenu.

c. Mode de synchronisation : choisissez la manière dont vous souhaitez mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous

synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation.

- Synchronisation complète : Indexez tout le contenu fraîchement, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
- Nouvelle synchronisation modifiée : Indexez uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
- Synchronisation nouvelle, modifiée ou supprimée : Indexez uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

d. Dans Synchroniser le calendrier d'exécution, pour Fréquence : à quelle fréquence Amazon Kendra sera synchronisée avec votre source de données.

e. Choisissez Suivant.

8. Sur la page Définir les mappages de champs, entrez les informations suivantes :

a. Sélectionnez parmi les champs de source de données par défaut générés (Document IDs, Titres des documents et Source URLs) que vous souhaitez Amazon Kendra mapper à l'index.

b. Ajouter un champ —Pour ajouter des champs de source de données personnalisés afin de créer un nom de champ d'index à mapper et le type de données du champ.

c. Choisissez Suivant.

9. Sur la page Réviser et créer, vérifiez que les informations saisies sont correctes, puis sélectionnez Ajouter une source de données. Vous pouvez également choisir de modifier vos informations depuis cette page. Votre source de données apparaîtra sur la page Sources de données une fois que la source de données aura été ajoutée avec succès.

API

Pour vous connecter Amazon Kendra à Oracle Database

Vous devez spécifier les éléments suivants à l'aide de l'TemplateConfigurationAPI :

- Source de données —Spécifiez le type de source de données tel que `JDBC` lorsque vous utilisez le schéma TemplateConfigurationJSON. Spécifiez également la source de données `TEMPLATE` lorsque vous appelez l'CreateDataSourceAPI.
- Type de base de données —Vous devez spécifier le type de base de données sous `oracle` la forme.
- Requête SQL : spécifiez des instructions de requête SQL telles que les opérations SELECT et JOIN. La taille des requêtes SQL doit être inférieure à 32 Ko. Amazon Kendra analysera tout le contenu de la base de données correspondant à votre requête.
- Mode de synchronisation : spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation. Vous pouvez choisir entre :
  - `FORCED_FULL_CRAWL`pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
  - `FULL_CRAWL`pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
  - `CHANGE_LOG`pour indexer uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
- Nom de ressource Amazon (ARN) secret : indiquez le nom de ressource Amazon (ARN) d'un Secrets Manager secret contenant les informations d'authentification que vous avez créées dans votre Oracle Database compte. Le secret est stocké dans une structure JSON avec les clés suivantes :

```
{
    "user name": "database user name",
```

```
    "password": "password"
}
```

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

- IAM role —Spécifiez à quel `RoleArn` moment vous appelez `CreateDataSource` pour fournir à un IAM rôle les autorisations d'accéder à votre Secrets Manager secret et d'appeler le public requis APIs pour le Oracle Database connecteur et Amazon Kendra. Pour plus d'informations, consultez la section IAM Rôles des sources de Oracle Database données.

Vous pouvez également ajouter les fonctionnalités optionnelles suivantes :

- Virtual Private Cloud (VPC) **`VpcConfiguration`** —Spécifiez le moment de votre appel. `CreateDataSource` Pour de plus amples informations, veuillez consulter Configuration Amazon Kendra pour utiliser un Amazon VPC.
- Filtres d'inclusion et d'exclusion : vous pouvez spécifier si vous souhaitez inclure un contenu spécifique en utilisant l'utilisateur IDs, les groupes, la source URLs, les horodatages et les fuseaux horaires.
- Filtrage du contexte utilisateur et contrôle d'accèsAmazon Kendra : analyse la liste de contrôle d'accès (ACL) de vos documents, si vous disposez d'une ACL pour vos documents. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.
- Mappages de champs : choisissez de mapper les champs de votre source de Oracle Database données à vos champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.

> **Note**
>
> Le champ du corps du document ou l'équivalent du corps du document pour vos documents est requis pour Amazon Kendra effectuer une recherche dans vos

documents. Vous devez associer le nom du champ du corps du document dans votre source de données au nom du champ d'index`_document_body`. Tous les autres champs sont facultatifs.

Pour obtenir la liste des autres clés JSON importantes à configurer, consultezSchéma du modèle de base de données Oracle.

Remarques

- Les lignes de base de données supprimées ne seront pas suivies lors de la Amazon Kendra vérification du contenu mis à jour.
- La taille des noms et des valeurs des champs d'une ligne de votre base de données ne peut pas dépasser 400 Ko.
- Si la source de données de votre base de données contient une grande quantité de données et que vous ne souhaitez pas Amazon Kendra indexer tout le contenu de votre base de données après la première synchronisation, vous pouvez choisir de ne synchroniser que les documents nouveaux, modifiés ou supprimés.
- Il est recommandé de fournir des informations d'identification Amazon Kendra de base de données en lecture seule.
- Il est recommandé d'éviter d'ajouter des tableaux contenant des données sensibles ou des informations personnelles identifiables (PII).

PostgreSQL

> **Note**
>
> PostgreSQLLe connecteur reste entièrement pris en charge pour les clients existants jusqu'au 31 mai 2026. Bien que ce connecteur ne soit plus disponible pour les nouveaux utilisateurs, les utilisateurs actuels peuvent continuer à l'utiliser sans interruption. Nous faisons constamment évoluer notre portefeuille de connecteurs afin de proposer des solutions plus évolutives et personnalisables. Pour les intégrations futures, nous vous recommandons d'explorer le framework de connecteurs personnalisés Amazon Kendra [1], conçu pour prendre en charge un plus large éventail de cas d'utilisation en entreprise avec une flexibilité accrue.

PostgreSQLest un système de gestion de base de données open source. Si vous êtes un PostgreSQL utilisateur, vous pouvez l'utiliser Amazon Kendra pour indexer votre source de PostgreSQL données. Le connecteur de source de Amazon Kendra PostgreSQL données prend en charge PostgreSQL 9.6.

Vous pouvez vous connecter Amazon Kendra à votre source de PostgreSQL données à l'aide de la Amazon Kendra console et de l'TemplateConfigurationAPI.

Pour résoudre les problèmes liés à votre connecteur de source de Amazon Kendra PostgreSQL données, consultezDépannage des sources de données.

Rubriques

- Fonctionnalités prises en charge
- Prérequis
- Instructions de connexion
- Remarques

Fonctionnalités prises en charge

- Mappages de champs
- Filtrage du contexte utilisateur
- Filtres d'inclusion/exclusion
- Synchronisation complète et incrémentielle du contenu
- Cloud privé virtuel (VPC)

Prérequis

Avant de pouvoir Amazon Kendra indexer votre source de PostgreSQL données, apportez ces modifications à vos AWS comptes PostgreSQL and.

DansPostgreSQL, assurez-vous d'avoir :

- Notez le nom d'utilisateur et le mot de passe de votre base de données.

> **Important**
>
> Il est recommandé de fournir des informations d'identification Amazon Kendra de base de données en lecture seule.

- Vous avez copié l'URL, le port et l'instance de votre hôte de base de données.
- Il est vérifié que chaque document est unique dans PostgreSQL et entre les autres sources de données que vous prévoyez d'utiliser pour le même index. Chaque source de données que vous souhaitez utiliser pour un index ne doit pas contenir le même document dans toutes les sources de données. IDs Les documents sont globaux par rapport à un index et doivent être uniques par index.

Dans votre Compte AWS, assurez-vous d'avoir :

- Création d'un Amazon Kendra index et, si vous utilisez l'API, notez l'ID de l'index.
- Vous avez créé un IAM rôle pour votre source de données et, si vous utilisez l'API, notez l'ARN du IAM rôle.

> **Note**
>
> Si vous modifiez votre type d'authentification et vos informations d'identification, vous devez mettre à jour votre IAM rôle pour accéder au bon identifiant AWS Secrets Manager secret.

- Stockez vos informations d'PostgreSQLauthentification dans un AWS Secrets Manager secret et, si vous utilisez l'API, notez l'ARN du secret.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

Si vous n'avez pas de IAM rôle ou de secret existant, vous pouvez utiliser la console pour créer un nouveau IAM rôle et un nouveau Secrets Manager secret lorsque vous connectez votre source de PostgreSQL données à Amazon Kendra. Si vous utilisez l'API, vous devez fournir l'ARN d'un IAM rôle et d'un Secrets Manager secret existants, ainsi qu'un identifiant d'index.

Instructions de connexion

Pour vous connecter Amazon Kendra à votre source de PostgreSQL données, vous devez fournir les détails de vos PostgreSQL informations d'identification afin de Amazon Kendra pouvoir accéder à vos données. Si vous n'avez pas encore configuré PostgreSQL pour Amazon Kendra voirPrérequis.

Console

Pour vous connecter Amazon Kendra à PostgreSQL

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la.
2. Dans le volet de navigation de gauche, choisissez Index, puis choisissez l'index que vous souhaitez utiliser dans la liste des index.

   > **Note**
   >
   > Vous pouvez choisir de configurer ou de modifier vos paramètres de contrôle d'accès utilisateur dans les paramètres de l'index.

3. Sur la page de démarrage, choisissez Ajouter une source de données.
4. Sur la page Ajouter une source de données, choisissez PostgreSQLconnecteur, puis sélectionnez Ajouter un connecteur. Si vous utilisez la version 2 (le cas échéant), choisissez le PostgreSQLconnecteur avec la balise « V2.0 ».
5. Sur la page Spécifier les détails de la source de données, entrez les informations suivantes :

   a. Dans Nom et description, pour Nom de la source de données : entrez le nom de votre source de données. Vous pouvez inclure des traits d'union, mais pas des espaces.
   b. (Facultatif) Description : entrez une description facultative pour votre source de données.
   c. Dans la langue par défaut : choisissez une langue pour filtrer vos documents pour l'index. Sauf indication contraire, la langue par défaut est l'anglais. La langue spécifiée dans les métadonnées du document remplace la langue sélectionnée.
   d. Dans Balises, pour Ajouter une nouvelle balise : incluez des balises facultatives pour rechercher et filtrer vos ressources ou suivre vos AWS coûts.

e. Choisissez Suivant.

6. Sur la page Définir l'accès et la sécurité, entrez les informations suivantes :

   a. Dans Source, entrez les informations suivantes :

   b. Hôte : entrez le nom d'hôte de la base de données.

   c. Port — Entrez le port de base de données.

   d. Instance — Entrez l'instance de base de données.

   e. Activer l'emplacement du certificat SSL : choisissez d'entrer le Amazon S3 chemin d'accès à votre fichier de certificat SSL.

   f. Dans Authentification, entrez les informations suivantes :

      - AWS Secrets Manager secret —Choisissez un secret existant ou créez-en un nouveau Secrets Manager pour stocker vos informations PostgreSQL d'authentification. Si vous choisissez de créer un nouveau secret, une fenêtre AWS Secrets Manager secrète s'ouvre.

        A. Entrez les informations suivantes dans la fenêtre Créer un AWS Secrets Manager secret :

           I. Nom secret : le nom de votre secret. Le préfixe « AmazonKendra - PostgreSQL - » est automatiquement ajouté à votre nom secret.

           II. Pour le nom d'utilisateur et le mot de passe de la base de données, entrez les valeurs d'identification d'authentification que vous avez copiées depuis votre base de données.

        B. Choisissez Enregistrer.

   g. Virtual Private Cloud (VPC) —Vous pouvez choisir d'utiliser un VPC. Dans ce cas, vous devez ajouter des sous-réseaux et des groupes de sécurité VPC.

   h. IAM rôle —Choisissez un IAM rôle existant ou créez-en un nouveau IAM pour accéder aux informations d'identification de votre référentiel et indexer le contenu.

> **Note**
>
> IAM les rôles utilisés pour les index ne peuvent pas être utilisés pour les sources de données. Si vous ne savez pas si un rôle existant est utilisé pour un index ou une FAQ, choisissez Créer un nouveau rôle pour éviter les erreurs.

i. Choisissez Suivant.

7. Sur la page Configurer les paramètres de synchronisation, entrez les informations suivantes :

   a. Dans le champ d'application de la synchronisation, choisissez l'une des options suivantes :

      - Requête SQL —Entrez des instructions de requête SQL telles que les opérations SELECT et JOIN. La taille des requêtes SQL doit être inférieure à 32 Ko. Amazon Kendra analysera tout le contenu de la base de données correspondant à votre requête.
      - Colonne clé primaire : indiquez la clé primaire pour la table de base de données. Cela identifie une table au sein de votre base de données.
      - Colonne de titre : indiquez le nom de la colonne de titre du document dans votre table de base de données.
      - Colonne du corps : indiquez le nom de la colonne du corps du document dans votre table de base de données.

   b. Dans Configuration supplémentaire — facultatif, choisissez l'une des options suivantes pour synchroniser un contenu spécifique au lieu de synchroniser tous les fichiers :

      - Colonnes détectant les modifications : entrez le nom des colonnes qui Amazon Kendra seront utilisées pour détecter les modifications de contenu. Amazon Kendra réindexera le contenu en cas de modification de l'une de ces colonnes.
      - IDsColonne des utilisateurs : entrez le nom de la colonne contenant l'utilisateur devant être autorisé IDs à accéder au contenu.
      - Colonne Groupes : entrez le nom de la colonne contenant les groupes autorisés à accéder au contenu.
      - URLsColonne source —Entrez le nom de la colonne contenant la source URLs à indexer.
      - Colonne horodatage : entrez le nom de la colonne contenant les horodatages. Amazon Kendra utilise les informations d'horodatage pour détecter les modifications apportées à votre contenu et synchroniser uniquement le contenu modifié.
      - Colonne fuseaux horaires : entrez le nom de la colonne contenant les fuseaux horaires du contenu à analyser.
      - Format d'horodatage : entrez le nom de la colonne contenant les formats d'horodatage à utiliser pour détecter les modifications de contenu et resynchroniser votre contenu.

c. Mode de synchronisation : choisissez la manière dont vous souhaitez mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation.

   - Synchronisation complète : Indexez tout le contenu fraîchement, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
   - Nouvelle synchronisation modifiée : Indexez uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
   - Synchronisation nouvelle, modifiée ou supprimée : Indexez uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

d. Dans Synchroniser le calendrier d'exécution, pour Fréquence : à quelle fréquence Amazon Kendra sera synchronisée avec votre source de données.

e. Choisissez Suivant.

8. Sur la page Définir les mappages de champs, entrez les informations suivantes :

   a. Sélectionnez parmi les champs de source de données par défaut générés (Document IDs, Titres des documents et Source URLs) que vous souhaitez Amazon Kendra mapper à l'index.

   b. Ajouter un champ —Pour ajouter des champs de source de données personnalisés afin de créer un nom de champ d'index à mapper et le type de données du champ.

   c. Choisissez Suivant.

9. Sur la page Réviser et créer, vérifiez que les informations saisies sont correctes, puis sélectionnez Ajouter une source de données. Vous pouvez également choisir de modifier vos informations à partir de cette page. Votre source de données apparaîtra sur la page Sources de données une fois que la source de données aura été ajoutée avec succès.

API

Pour vous connecter Amazon Kendra à PostgreSQL

Vous devez spécifier les éléments suivants à l'aide de l'TemplateConfigurationAPI :

- Source de données —Spécifiez le type de source de données tel que `JDBC` lorsque vous utilisez le schéma TemplateConfigurationJSON. Spécifiez également la source de données `TEMPLATE` lorsque vous appelez l'CreateDataSourceAPI.
- Type de base de données —Vous devez spécifier le type de base de données sous `postgresql` la forme.
- Requête SQL : spécifiez des instructions de requête SQL telles que les opérations SELECT et JOIN. La taille des requêtes SQL doit être inférieure à 32 Ko. Amazon Kendra analysera tout le contenu de la base de données correspondant à votre requête.
- Mode de synchronisation : spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation. Vous pouvez choisir entre :
  - `FORCED_FULL_CRAWL`pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
  - `FULL_CRAWL`pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
  - `CHANGE_LOG`pour indexer uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
- Nom de ressource Amazon (ARN) secret : indiquez le nom de ressource Amazon (ARN) d'un Secrets Manager secret contenant les informations d'authentification que vous avez créées dans votre PostgreSQL compte. Le secret est stocké dans une structure JSON avec les clés suivantes :

```
{
    "user name": "database user name",
```

```
    "password": "password"
}
```

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

- IAM role —Spécifiez à quel `RoleArn` moment vous appelez `CreateDataSource` pour fournir à un IAM rôle les autorisations d'accéder à votre Secrets Manager secret et d'appeler le public requis APIs pour le PostgreSQL connecteur et Amazon Kendra. Pour plus d'informations, consultez la section IAM Rôles des sources de PostgreSQL données.

Vous pouvez également ajouter les fonctionnalités optionnelles suivantes :

- Virtual Private Cloud (VPC) **`VpcConfiguration`** —Spécifiez le moment de votre appel. `CreateDataSource` Pour de plus amples informations, veuillez consulter Configuration Amazon Kendra pour utiliser un Amazon VPC.
- Filtres d'inclusion et d'exclusion : vous pouvez spécifier si vous souhaitez inclure un contenu spécifique en utilisant l'utilisateur IDs, les groupes, la source URLs, les horodatages et les fuseaux horaires.
- Filtrage du contexte utilisateur et contrôle d'accèsAmazon Kendra : analyse la liste de contrôle d'accès (ACL) de vos documents, si vous disposez d'une ACL pour vos documents. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.
- Mappages de champs : choisissez de mapper les champs de votre source de PostgreSQL données à vos champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.

> **Note**
>
> Le champ du corps du document ou l'équivalent du corps du document pour vos documents est requis pour Amazon Kendra effectuer une recherche dans vos

documents. Vous devez associer le nom du champ du corps du document dans votre source de données au nom du champ d'index`_document_body`. Tous les autres champs sont facultatifs.

Pour obtenir la liste des autres clés JSON importantes à configurer, consultezSchéma du modèle PostgreSQL.

Remarques

- Les lignes de base de données supprimées ne seront pas suivies lors de la Amazon Kendra vérification du contenu mis à jour.
- La taille des noms et des valeurs des champs d'une ligne de votre base de données ne peut pas dépasser 400 Ko.
- Si la source de données de votre base de données contient une grande quantité de données et que vous ne souhaitez pas Amazon Kendra indexer tout le contenu de votre base de données après la première synchronisation, vous pouvez choisir de ne synchroniser que les documents nouveaux, modifiés ou supprimés.
- Il est recommandé de fournir des informations d'identification Amazon Kendra de base de données en lecture seule.
- Il est recommandé d'éviter d'ajouter des tableaux contenant des données sensibles ou des informations personnelles identifiables (PII).

Quip

Note

Le connecteur Quip reste entièrement compatible pour les clients existants jusqu'au 31 mai 2026. Bien que ce connecteur ne soit plus disponible pour les nouveaux utilisateurs, les utilisateurs actuels peuvent continuer à l'utiliser sans interruption. Nous faisons constamment évoluer notre portefeuille de connecteurs afin de proposer des solutions plus évolutives et personnalisables. Pour les intégrations futures, nous vous recommandons d'explorer le framework de connecteurs personnalisés Amazon Kendra [1], conçu pour prendre en charge un plus large éventail de cas d'utilisation en entreprise avec une flexibilité accrue.

Quip est un logiciel de productivité collaboratif qui offre des fonctionnalités de création de documents en temps réel. Vous pouvez l'utiliser Amazon Kendra pour indexer vos dossiers, fichiers, commentaires de fichiers, forums de discussion et pièces jointes Quip.

Vous pouvez vous connecter Amazon Kendra à votre source de données Quip à l'aide de la Amazon Kendra console et de l'QuipConfigurationAPI.

Pour résoudre les problèmes liés à votre connecteur de source de données Amazon Kendra Quip, consultezDépannage des sources de données.

Rubriques

- Fonctionnalités prises en charge
- Prérequis
- Instructions de connexion
- En savoir plus

Fonctionnalités prises en charge

Amazon Kendra Le connecteur de source de données Quip prend en charge les fonctionnalités suivantes :

- Mappages de champs
- Contrôle d'accès utilisateur
- Filtres d'inclusion/exclusion
- Cloud privé virtuel (VPC)

Prérequis

Avant de pouvoir Amazon Kendra indexer votre source de données Quip, apportez ces modifications à votre Quip et AWS à vos comptes.

Dans Quip, assurez-vous d'avoir :

- Un compte Quip doté d'autorisations administratives.
- Vous avez créé des identifiants d'authentification Quip qui incluent un jeton d'accès personnel. Le jeton est utilisé comme identifiant d'authentification stocké dans un AWS Secrets Manager secret. Consultez la documentation Quip sur l'authentification pour plus d'informations.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

- Vous avez copié le domaine de votre site Quip. Par exemple, *`https://quip-company.quipdomain.com/browse`* où se *`quipdomain`* trouve le domaine ?
- Il est vérifié que chaque document est unique dans Quip et dans les autres sources de données que vous prévoyez d'utiliser pour le même index. Chaque source de données que vous souhaitez utiliser pour un index ne doit pas contenir le même document dans toutes les sources de données. IDs Les documents sont globaux par rapport à un index et doivent être uniques par index.

Dans votre Compte AWS, assurez-vous d'avoir :

- Création d'un Amazon Kendra index et, si vous utilisez l'API, notez l'ID de l'index.
- Vous avez créé un IAM rôle pour votre source de données et, si vous utilisez l'API, notez l'ARN du IAM rôle.

> **Note**
>
> Si vous modifiez votre type d'authentification et vos informations d'identification, vous devez mettre à jour votre IAM rôle pour accéder au bon identifiant AWS Secrets Manager secret.

- Stockez vos informations d'authentification Quip dans un AWS Secrets Manager secret et, si vous utilisez l'API, notez l'ARN du secret.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations

d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

Si vous n'avez pas de IAM rôle ou de secret existant, vous pouvez utiliser la console pour créer un nouveau IAM rôle et un nouveau Secrets Manager secret lorsque vous connectez votre source de données Quip à Amazon Kendra. Si vous utilisez l'API, vous devez fournir l'ARN d'un IAM rôle et d'un Secrets Manager secret existants, ainsi qu'un identifiant d'index.

Instructions de connexion

Pour vous connecter Amazon Kendra à votre source de données Quip, vous devez fournir les détails nécessaires de votre source de données Quip afin de Amazon Kendra pouvoir accéder à vos données. Si vous n'avez pas encore configuré Quip pour Amazon Kendra, consultezPrérequis.

Console

Pour vous connecter Amazon Kendra à Quip

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la.
2. Dans le volet de navigation de gauche, choisissez Index, puis choisissez l'index que vous souhaitez utiliser dans la liste des index.

   > **Note**
   >
   > Vous pouvez choisir de configurer ou de modifier vos paramètres de contrôle d'accès utilisateur dans les paramètres de l'index.

3. Sur la page de démarrage, choisissez Ajouter une source de données.
4. Sur la page Ajouter une source de données, choisissez le connecteur Quip, puis sélectionnez Ajouter un connecteur. Si vous utilisez la version 2 (le cas échéant), choisissez le connecteur Quip avec le tag « V2.0 ».
5. Sur la page Spécifier les détails de la source de données, entrez les informations suivantes :

   a. Dans Nom et description, pour Nom de la source de données : entrez le nom de votre source de données. Vous pouvez inclure des tirets, mais pas des espaces.

   b. (Facultatif) Description : entrez une description facultative pour votre source de données.

c. Dans la langue par défaut : choisissez une langue pour filtrer vos documents pour l'index. Sauf indication contraire, la langue par défaut est l'anglais. La langue spécifiée dans les métadonnées du document remplace la langue sélectionnée.

d. Dans Balises, pour Ajouter une nouvelle balise : incluez des balises facultatives pour rechercher et filtrer vos ressources ou suivre vos AWS coûts.

e. Choisissez Suivant.

6. Sur la page Définir l'accès et la sécurité, entrez les informations suivantes :

a. Nom de domaine Quip —Entrez le nom de domaine Quip que vous avez copié depuis votre compte Quip.

b. AWS Secrets Manager secret —Choisissez un secret existant ou créez-en un nouveau Secrets Manager pour stocker vos informations d'authentification Quip. Si vous choisissez de créer un nouveau secret, une fenêtre AWS Secrets Manager secrète s'ouvre.

i. Entrez les informations suivantes dans la fenêtre Créer un AWS Secrets Manager secret :

A. Nom secret : le nom de votre secret. Le préfixe « AmazonKendra -Quip » est automatiquement ajouté à votre nom secret.

B. Jeton Quip —Entrez l'accès personnel Quip configuré pour Quip.

ii. Ajoutez et enregistrez votre secret.

c. Virtual Private Cloud (VPC) —Vous pouvez choisir d'utiliser un VPC. Dans ce cas, vous devez ajouter des sous-réseaux et des groupes de sécurité VPC.

d. IAM rôle —Choisissez un IAM rôle existant ou créez-en un nouveau IAM pour accéder aux informations d'identification de votre référentiel et indexer le contenu.

> **Note**
>
> IAM les rôles utilisés pour les index ne peuvent pas être utilisés pour les sources de données. Si vous ne savez pas si un rôle existant est utilisé pour un index ou une FAQ, choisissez Créer un nouveau rôle pour éviter les erreurs.

e. Choisissez Suivant.

7. Sur la page Configurer les paramètres de synchronisation, entrez les informations suivantes :

a. Ajouter le dossier Quip IDs à explorer : le dossier Quip que IDs vous souhaitez explorer.

> **Note**
>
> Pour explorer un dossier racine, y compris tous les sous-dossiers et documents qu'il contient, ajoutez l'ID du dossier racine. Pour explorer des sous-dossiers spécifiques, ajoutez-les. IDs

b. Configuration supplémentaire (types de contenu) : entrez les types de contenu que vous souhaitez analyser.

c. Modèles Regex : modèles d'expressions régulières permettant d'inclure ou d'exclure certains fichiers. Vous pouvez ajouter jusqu'à 100 motifs.

d. Dans Synchroniser le calendrier d'exécution, pour Fréquence : choisissez la fréquence à laquelle vous souhaitez synchroniser le contenu de votre source de données et mettre à jour votre index

e. Choisissez Suivant.

8. Sur la page Définir les mappages de champs, entrez les informations suivantes :

a. Sélectionnez parmi les champs de source de données par défaut générés que vous souhaitez Amazon Kendra mapper à l'index.

b. Ajouter un champ —Pour ajouter des champs de source de données personnalisés afin de créer un nom de champ d'index à mapper et le type de données du champ.

c. Choisissez Suivant.

9. Sur la page Réviser et créer, vérifiez que les informations saisies sont correctes, puis sélectionnez Ajouter une source de données. Vous pouvez également choisir de modifier vos informations depuis cette page. Votre source de données apparaîtra sur la page Sources de données une fois que la source de données aura été ajoutée avec succès.

API

Pour vous connecter Amazon Kendra à Quip

Vous devez spécifier les éléments suivants à l'aide de l'QuipConfiguration API :

- Domaine du site Quip : par exemple, *`https://quip-company.quipdomain.com/browse`* où se *`quipdomain`* trouve le domaine ?

- Nom de ressource Amazon (ARN) secret : indiquez le nom de ressource Amazon (ARN) d'un Secrets Manager secret contenant les informations d'authentification de votre compte Quip. Le secret est stocké dans une structure JSON avec les clés suivantes :

```
{
    "accessToken": "token"
}
```

- IAM role —Spécifiez à quel `RoleArn` moment vous appelez `CreateDataSource` pour fournir à un IAM rôle les autorisations d'accéder à votre Secrets Manager secret et pour appeler le public requis APIs pour le connecteur Quip et. Amazon Kendra Pour plus d'informations, consultez la section IAM Rôles des sources de données Quip.

Vous pouvez également ajouter les fonctionnalités optionnelles suivantes :

- Virtual Private Cloud (VPC) —Spécifiez dans le `VpcConfiguration` cadre de la configuration de la source de données. Consultez la section Configuration Amazon Kendra pour utiliser un VPC.
- Filtres d'inclusion et d'exclusion : spécifiez si vous souhaitez inclure ou exclure certains fichiers.

> **Note**
>
> La plupart des sources de données utilisent des modèles d'expressions régulières, qui sont des modèles d'inclusion ou d'exclusion appelés filtres. Si vous spécifiez un filtre d'inclusion, seul le contenu correspondant au filtre d'inclusion est indexé. Tout document qui ne correspond pas au filtre d'inclusion n'est pas indexé. Si vous spécifiez un filtre d'inclusion et d'exclusion, les documents correspondant au filtre d'exclusion ne sont pas indexés, même s'ils correspondent au filtre d'inclusion.

- Dossiers : spécifiez les dossiers et sous-dossiers Quip que vous souhaitez indexer

> **Note**
>
> Pour explorer un dossier racine, y compris tous les sous-dossiers et documents qu'il contient, entrez l'ID du dossier racine. Pour explorer des sous-dossiers spécifiques, ajoutez-les. IDs

- Pièces jointes, salons de discussion, commentaires sur les fichiers : choisissez d'inclure ou non l'exploration des pièces jointes, du contenu des salons de discussion et des commentaires sur les fichiers.
- Filtrage du contexte utilisateur et contrôle d'accèsAmazon Kendra : analyse la liste de contrôle d'accès (ACL) de vos documents, si vous disposez d'une ACL pour vos documents. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.
- Mappages de champs : choisissez de mapper les champs de votre source de données Quip à vos Amazon Kendra champs d'index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.

> **Note**
>
> Le champ du corps du document ou l'équivalent du corps du document pour vos documents est requis pour Amazon Kendra effectuer une recherche dans vos documents. Vous devez associer le nom du champ du corps du document dans votre source de données au nom du champ d'index`_document_body`. Tous les autres champs sont facultatifs.

En savoir plus

Pour en savoir plus sur l'intégration Amazon Kendra à votre source de données Quip, consultez :

- Recherchez des connaissances dans des documents Quip grâce à la recherche intelligente à l'aide du connecteur Quip pour Amazon Kendra

Salesforce

Salesforce est un outil de gestion de la relation client (CRM) permettant de gérer les équipes de support, de vente et de marketing. Vous pouvez l'utiliser Amazon Kendra pour indexer vos objets standard Salesforce et même vos objets personnalisés.

Vous pouvez vous connecter Amazon Kendra à votre source de données Salesforce à l'aide de la Amazon Kendra console, de l'TemplateConfigurationAPI ou de l'SalesforceConfigurationAPI.

Amazon Kendra possède deux versions du connecteur Salesforce. Les fonctionnalités prises en charge par chaque version incluent :

Connecteur Salesforce V1.0/API SalesforceConfiguration

- Mappages de champs
- Contrôle d'accès utilisateur
- Filtres d'inclusion/exclusion

Connecteur Salesforce V2.0/API TemplateConfiguration

- Mappages de champs
- Contrôle d'accès utilisateur
- Filtres d'inclusion/exclusion
- Synchronisation complète et incrémentielle du contenu
- Cloud privé virtuel (VPC)

> **Note**
>
> Le connecteur Salesforce SalesforceConfiguration V1.0/API a pris fin en 2023. Nous vous recommandons de migrer vers ou d'utiliser le connecteur Salesforce TemplateConfiguration V2.0/API.

Pour résoudre les problèmes liés à votre connecteur de source de données Amazon Kendra Salesforce, consultezDépannage des sources de données.

Rubriques

- Connecteur Salesforce V1.0
- Connecteur Salesforce V2.0

Connecteur Salesforce V1.0

Salesforce est un outil de gestion de la relation client (CRM) permettant de gérer les équipes de support, de vente et de marketing. Vous pouvez l'utiliser Amazon Kendra pour indexer vos objets standard Salesforce et même vos objets personnalisés.

> **Important**
>
> Amazon Kendra utilise la version 48 de l'API Salesforce. L'API Salesforce limite le nombre de demandes que vous pouvez effectuer par jour. Si Salesforce dépasse le nombre de demandes, il réessaie jusqu'à ce qu'il soit en mesure de continuer.

> **Note**
>
> Le connecteur Salesforce SalesforceConfiguration V1.0/API a pris fin en 2023. Nous vous recommandons de migrer vers ou d'utiliser le connecteur Salesforce TemplateConfiguration V2.0/API.

Pour résoudre les problèmes liés à votre connecteur de source de données Amazon Kendra Salesforce, consultezDépannage des sources de données.

Rubriques

- Fonctionnalités prises en charge
- Prérequis
- Instructions de connexion

Fonctionnalités prises en charge

Amazon Kendra Le connecteur de source de données Salesforce prend en charge les fonctionnalités suivantes :

- Mappages de champs
- Contrôle d'accès utilisateur
- Filtres d'inclusion/exclusion

Prérequis

Avant de pouvoir l'utiliser Amazon Kendra pour indexer votre source de données Salesforce, apportez ces modifications à votre Salesforce et à vos AWS comptes.

Dans Salesforce, assurez-vous que vous disposez des éléments suivants :

- J'ai créé un compte Salesforce et j'ai noté le nom d'utilisateur et le mot de passe que vous utilisez pour vous connecter à Salesforce.
- Vous avez créé un compte Salesforce Connected App en OAuth activant et en copiant la clé client (ID client) et le secret client (secret client) attribués à votre application Salesforce Connected. L'ID client et le secret du client sont utilisés comme informations d'authentification stockées dans un AWS Secrets Manager secret. Consultez la documentation Salesforce sur les applications connectées pour plus d'informations.

  > **Note**
  >
  > Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

- Le jeton de sécurité Salesforce associé au compte utilisé pour se connecter à Salesforce a été copié.
- Vous avez copié l'URL de l'instance Salesforce que vous souhaitez indexer. Il s'agit généralement de `https://<company>.salesforce.com/`. Le serveur doit exécuter une application connectée à Salesforce.
- Vous avez ajouté des informations d'identification à votre serveur Salesforce pour un utilisateur ayant un accès en lecture seule à Salesforce en clonant le ReadOnly profil, puis en ajoutant les autorisations Afficher toutes les données et Gérer les articles. Ces informations d'identification identifient l'utilisateur qui établit la connexion et l'application connectée Salesforce à laquelle Amazon Kendra elle se connecte.
- Il est vérifié que chaque document est unique dans Salesforce et dans les autres sources de données que vous prévoyez d'utiliser pour le même index. Chaque source de données que vous souhaitez utiliser pour un index ne doit pas contenir le même document dans toutes les sources de données. IDs Les documents sont globaux par rapport à un index et doivent être uniques par index.

Dans votre Compte AWS, assurez-vous d'avoir :

- Création d'un Amazon Kendra index et, si vous utilisez l'API, notez l'ID de l'index.
- Vous avez créé un IAM rôle pour votre source de données et, si vous utilisez l'API, notez l'ARN du IAM rôle.

> **Note**
>
> Si vous modifiez votre type d'authentification et vos informations d'identification, vous devez mettre à jour votre IAM rôle pour accéder au bon identifiant AWS Secrets Manager secret.

- Stockez vos informations d'authentification Salesforce dans un AWS Secrets Manager secret et, si vous utilisez l'API, notez l'ARN du secret.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

Si vous n'avez pas de IAM rôle ou de secret existant, vous pouvez utiliser la console pour créer un nouveau IAM rôle et un nouveau Secrets Manager secret lorsque vous connectez votre source de données Salesforce à Amazon Kendra. Si vous utilisez l'API, vous devez fournir l'ARN d'un IAM rôle et d'un Secrets Manager secret existants, ainsi qu'un identifiant d'index.

Instructions de connexion

Pour vous connecter Amazon Kendra à votre source de données Salesforce, vous devez fournir les informations nécessaires sur votre source de données Salesforce afin de Amazon Kendra pouvoir accéder à vos données. Si vous n'avez pas encore configuré Salesforce pour Amazon Kendra voirPrérequis.

Console

Pour vous connecter Amazon Kendra à Salesforce

1. Connectez-vous à la console AWS de gestion et Amazon Kendra ouvrez-la.
2. Dans le volet de navigation de gauche, choisissez Index, puis choisissez l'index que vous souhaitez utiliser dans la liste des index.

> **Note**
>
> Vous pouvez choisir de configurer ou de modifier vos paramètres de contrôle d'accès utilisateur dans les paramètres de l'index.

3. Sur la page de démarrage, choisissez Ajouter une source de données.
4. Sur la page Ajouter une source de données, choisissez le connecteur Salesforce V1.0, puis sélectionnez Ajouter un connecteur.
5. Sur la page Spécifier les détails de la source de données, entrez les informations suivantes :

   a. Nom de la source de données —Entrez le nom de votre source de données. Vous pouvez inclure des traits d'union, mais pas des espaces.

   b. (Facultatif) Description : entrez une description facultative pour votre source de données.

   c. Langue par défaut : langue permettant de filtrer vos documents pour l'index. Sauf indication contraire, la langue par défaut est l'anglais. La langue spécifiée dans les métadonnées remplace la langue sélectionnée.

   d. Ajouter une nouvelle balise : des balises pour rechercher et filtrer vos ressources ou suivre vos coûts partagés.

   e. Choisissez Suivant.
6. Sur la page Définir l'accès et la sécurité, entrez les informations suivantes :

   a. URL Salesforce —Entrez l'URL de l'instance du site Salesforce que vous souhaitez indexer.

   b. Pour Type d'authentification, choisissez entre Existant et Nouveau pour stocker vos informations d'authentification Salesforce. Si vous choisissez de créer un nouveau secret, une fenêtre AWS Secrets Manager secrète s'ouvre.

      - Entrez les informations suivantes dans la fenêtre Créer un AWS Secrets Manager secret :

        A. Nom secret : le nom de votre secret. Le préfixe « AmazonKendra -Salesforce-» est automatiquement ajouté à votre nom secret.

        B. Pour le nom d'utilisateur, le mot de passe, le jeton de sécurité, la clé du client, le secret du consommateur et l'URL d'authentification, entrez les valeurs

d'identification d'authentification que vous avez créées dans votre compte Salesforce.

C. Choisissez Enregistrer l'authentification.

c. IAM rôle —Choisissez un IAM rôle existant ou créez-en un nouveau IAM pour accéder aux informations d'identification de votre référentiel et indexer le contenu.

> **Note**
>
> IAM les rôles utilisés pour les index ne peuvent pas être utilisés pour les sources de données. Si vous ne savez pas si un rôle existant est utilisé pour un index ou une FAQ, choisissez Créer un nouveau rôle pour éviter les erreurs.

d. Choisissez Suivant.

7. Sur la page Configurer les paramètres de synchronisation, entrez les informations suivantes :

a. Pour explorer les pièces jointes : sélectionnez cette option pour analyser tous les objets, articles et flux joints.

b. Pour les objets standard, les articles de connaissances et les fils Chatter, sélectionnez les entités Salesforce ou les types de contenu que vous souhaitez analyser.

> **Note**
>
> Vous devez fournir des informations de configuration pour indexer au moins un objet standard, un article de connaissances ou un fil de discussion.
> Si vous choisissez d'explorer les articles de connaissances, vous devez spécifier les types d'articles de connaissances à indexer, le nom des articles et indiquer si vous souhaitez indexer les champs standard de tous les articles de connaissances ou uniquement les champs d'un type d'article personnalisé. Si vous choisissez d'indexer des articles personnalisés, vous devez spécifier le nom interne du type d'article. Vous pouvez spécifier jusqu'à 10 types d'articles.

c. Fréquence : fréquence Amazon Kendra de synchronisation avec votre source de données.

d. Choisissez Suivant.

8. Sur la page Définir les mappages de champs, entrez les informations suivantes :

a. Pour l'article de connaissance standard, les pièces jointes aux objets standard et les mappages de champs suggérés supplémentaires, sélectionnez parmi les champs de source de données par défaut Amazon Kendra générés que vous souhaitez mapper à votre index.

> **Note**
>
> Un index mappé à `_document_body` est requis. Vous ne pouvez pas modifier le mappage entre le `Salesforce ID` champ et le Amazon Kendra `_document_id` champ.

b. Ajouter un champ —Pour ajouter des champs de source de données personnalisés afin de créer un nom de champ d'index à mapper et le type de données du champ.

c. Choisissez Suivant.

9. Sur la page Réviser et créer, vérifiez que les informations saisies sont correctes, puis sélectionnez Ajouter une source de données. Vous pouvez également choisir de modifier vos informations depuis cette page. Votre source de données apparaîtra sur la page Sources de données une fois que la source de données aura été ajoutée avec succès.

API

Pour vous connecter Amazon Kendra à Salesforce

Vous devez spécifier l'SalesforceConfigurationAPI suivante :

- URL du serveur : URL de l'instance du site Salesforce que vous souhaitez indexer.
- Nom de ressource Amazon secret (ARN) : indiquez le nom de ressource Amazon (ARN) d'un Secrets Manager secret contenant les informations d'authentification de votre compte Salesforce. Le secret est stocké dans une structure JSON avec les clés suivantes :

```
{
    "authenticationUrl": "OAUTH endpoint that Amazon Kendra connects to get an
 OAUTH token",
    "consumerKey": "Application public key generated when you created your
 Salesforce application",
    "consumerSecret": "Application private key generated when you created your
 Salesforce application.",
```

```
    "password": "Password associated with the user logging in to the Salesforce
 instance",
    "securityToken": "Token associated with the user account logging in to the
 Salesforce instance",
    "username": "User name of the user logging in to the Salesforce instance"
}
```

- IAM role —Spécifiez à quel `RoleArn` moment vous appelez `CreateDataSource` pour fournir à un IAM rôle les autorisations d'accéder à votre Secrets Manager secret et d'appeler le public requis APIs pour le connecteur Salesforce et Amazon Kendra. Pour plus d'informations, consultez la section IAM Rôles pour les sources de données Salesforce.
- Vous devez fournir des informations de configuration pour indexer au moins un objet standard, un article de connaissances ou un fil de discussion.
  - Objets standard : si vous choisissez d'analyser des objets standard, vous devez spécifier le nom de l'objet standard et le nom du champ dans la table des objets standard qui contient le contenu du document.
  - Articles de connaissances : si vous choisissez d'explorer les articles de connaissances, vous devez spécifier les types d'articles de connaissances à indexer, l'état des articles de connaissances à indexer et indiquer si vous souhaitez indexer les champs standard de tous les articles de connaissances ou uniquement les champs d'un type d'article personnalisé.
  - Flux Chatter : si vous choisissez d'analyser les flux Chatter, vous devez spécifier le nom de la colonne du FeedItem tableau Salesforce qui contient le contenu à indexer.

Vous pouvez également ajouter les fonctionnalités optionnelles suivantes :

- Filtres d'inclusion et d'exclusion : spécifiez si vous souhaitez inclure ou exclure certaines pièces jointes.

> **Note**
>
> La plupart des sources de données utilisent des modèles d'expressions régulières, qui sont des modèles d'inclusion ou d'exclusion appelés filtres. Si vous spécifiez un filtre d'inclusion, seul le contenu correspondant au filtre d'inclusion est indexé. Tout document qui ne correspond pas au filtre d'inclusion n'est pas indexé. Si vous spécifiez un filtre d'inclusion et d'exclusion, les documents correspondant au filtre d'exclusion ne sont pas indexés, même s'ils correspondent au filtre d'inclusion.

- Mappages de champs : choisissez de mapper les champs de votre source de données Salesforce à vos champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.

  > **Note**
  >
  > Le champ du corps du document ou l'équivalent du corps du document pour vos documents est requis pour Amazon Kendra effectuer une recherche dans vos documents. Vous devez associer le nom du champ du corps du document dans votre source de données au nom du champ d'index`_document_body`. Tous les autres champs sont facultatifs.

- Filtrage du contexte utilisateur et contrôle d'accèsAmazon Kendra : analyse la liste de contrôle d'accès (ACL) de vos documents, si vous disposez d'une ACL pour vos documents. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.

Connecteur Salesforce V2.0

Salesforce est un outil de gestion de la relation client (CRM) permettant de gérer les équipes de support, de vente et de marketing. Vous pouvez l'utiliser Amazon Kendra pour indexer vos objets standard Salesforce et même vos objets personnalisés.

Le connecteur de source de données Amazon Kendra Salesforce prend en charge les éditions Salesforce suivantes : Developer Edition et Enterprise Edition.

> **Note**
>
> Le connecteur Salesforce SalesforceConfiguration V1.0/API a pris fin en 2023. Nous vous recommandons de migrer vers ou d'utiliser le connecteur Salesforce TemplateConfiguration V2.0/API.

Pour résoudre les problèmes liés à votre connecteur de source de données Amazon Kendra Salesforce, consultezDépannage des sources de données.

Rubriques

- Fonctionnalités prises en charge
- Prérequis
- Instructions de connexion
- En savoir plus
- Remarques

Fonctionnalités prises en charge

Amazon Kendra Le connecteur de source de données Salesforce prend en charge les fonctionnalités suivantes :

- Mappages de champs
- Contrôle d'accès utilisateur
- Filtres d'inclusion/exclusion
- Synchronisation complète et incrémentielle du contenu
- Cloud privé virtuel (VPC)

Prérequis

Avant de pouvoir l'utiliser Amazon Kendra pour indexer votre source de données Salesforce, apportez ces modifications à votre Salesforce et à vos AWS comptes.

Dans Salesforce, assurez-vous que vous disposez des éléments suivants :

- J'ai créé un compte administratif Salesforce et j'ai noté le nom d'utilisateur et le mot de passe que vous utilisez pour vous connecter à Salesforce.
- Le jeton de sécurité Salesforce associé au compte utilisé pour se connecter à Salesforce a été copié.
- Vous avez créé un compte Salesforce Connected App en OAuth activant et en copiant la clé client (ID client) et le secret client (secret client) attribués à votre application Salesforce Connected. L'ID client et le secret du client sont utilisés comme informations d'authentification stockées dans un AWS Secrets Manager secret. Consultez la documentation Salesforce sur les applications connectées pour plus d'informations.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

- Vous avez copié l'URL de l'instance Salesforce que vous souhaitez indexer. Il s'agit généralement de *`https://<company>.salesforce.com/`*. Le serveur doit exécuter une application connectée à Salesforce.
- Vous avez ajouté des informations d'identification à votre serveur Salesforce pour un utilisateur ayant un accès en lecture seule à Salesforce en clonant le ReadOnly profil, puis en ajoutant les autorisations Afficher toutes les données et Gérer les articles. Ces informations d'identification identifient l'utilisateur qui établit la connexion et l'application connectée Salesforce à laquelle Amazon Kendra elle se connecte.
- Il est vérifié que chaque document est unique dans Salesforce et dans les autres sources de données que vous prévoyez d'utiliser pour le même index. Chaque source de données que vous souhaitez utiliser pour un index ne doit pas contenir le même document dans toutes les sources de données. IDs Les documents sont globaux par rapport à un index et doivent être uniques par index.

Dans votre Compte AWS, assurez-vous d'avoir :

- Création d'un Amazon Kendra index et, si vous utilisez l'API, notez l'ID de l'index.
- Vous avez créé un IAM rôle pour votre source de données et, si vous utilisez l'API, notez l'ARN du IAM rôle.

> **Note**
>
> Si vous modifiez votre type d'authentification et vos informations d'identification, vous devez mettre à jour votre IAM rôle pour accéder au bon identifiant AWS Secrets Manager secret.

- Stockez vos informations d'authentification Salesforce dans un AWS Secrets Manager secret et, si vous utilisez l'API, notez l'ARN du secret.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

Si vous n'avez pas de IAM rôle ou de secret existant, vous pouvez utiliser la console pour créer un nouveau IAM rôle et un nouveau Secrets Manager secret lorsque vous connectez votre source de données Salesforce à Amazon Kendra. Si vous utilisez l'API, vous devez fournir l'ARN d'un IAM rôle et d'un Secrets Manager secret existants, ainsi qu'un identifiant d'index.

Instructions de connexion

Pour vous connecter Amazon Kendra à votre source de données Salesforce, vous devez fournir les informations nécessaires sur votre source de données Salesforce afin de Amazon Kendra pouvoir accéder à vos données. Si vous n'avez pas encore configuré Salesforce pour Amazon Kendra voirPrérequis.

Console

Pour vous connecter Amazon Kendra à Salesforce :

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la.
2. Dans le volet de navigation de gauche, choisissez Index, puis choisissez l'index que vous souhaitez utiliser dans la liste des index.

   > **Note**
   >
   > Vous pouvez choisir de configurer ou de modifier vos paramètres de contrôle d'accès utilisateur dans les paramètres de l'index.

3. Sur la page de démarrage, choisissez Ajouter une source de données.
4. Sur la page Ajouter une source de données, choisissez le connecteur Salesforce, puis sélectionnez Ajouter un connecteur. Si vous utilisez la version 2 (le cas échéant), choisissez le connecteur Salesforce avec la balise « V2.0 ».

5. Sur la page Spécifier les détails de la source de données, entrez les informations suivantes :

   a. Dans Nom et description, pour Nom de la source de données : entrez le nom de votre source de données. Vous pouvez inclure des traits d'union, mais pas des espaces.

   b. (Facultatif) Description : entrez une description facultative pour votre source de données.

   c. Dans la langue par défaut : choisissez une langue pour filtrer vos documents pour l'index. Sauf indication contraire, la langue par défaut est l'anglais. La langue spécifiée dans les métadonnées du document remplace la langue sélectionnée.

   d. Dans Balises, pour Ajouter une nouvelle balise : incluez des balises facultatives pour rechercher et filtrer vos ressources ou suivre vos AWS coûts.

   e. Choisissez Suivant.

6. Sur la page Définir l'accès et la sécurité, entrez les informations suivantes :

   a. URL Salesforce —Entrez l'URL de l'instance du site Salesforce que vous souhaitez indexer.

   b. Autorisation : activez ou désactivez les informations de la liste de contrôle d'accès (ACL) pour vos documents, si vous disposez d'une ACL et que vous souhaitez l'utiliser pour le contrôle d'accès. L'ACL indique les documents auxquels les utilisateurs et les groupes peuvent accéder. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.

   c. Entrez un secret existant ou, si vous en créez un nouveau, une fenêtre de AWS Secrets Manager secret s'ouvre.

      - Authentification : entrez les informations suivantes dans la fenêtre Créer un AWS Secrets Manager secret :

        A. Nom secret : le nom de votre secret. Le préfixe « AmazonKendra -Salesforce-» est automatiquement ajouté à votre nom secret.

        B. Pour le nom d'utilisateur, le mot de passe, le jeton de sécurité, la clé du client, le secret du consommateur et l'URL d'authentification, entrez les valeurs d'identification d'authentification que vous avez générées et téléchargées depuis votre compte Salesforce.

> **Note**
>
> Si vous utilisez Salesforce Developer Edition, utilisez `https://login.salesforce.com/services/oauth2/token` l'URL de connexion My Domain (par exemple *`https://MyCompany.my.salesforce.com`*) comme URL d'authentification. Si vous utilisez Salesforce Sandbox Edition, utilisez `https://test.salesforce.com/services/oauth2/token` l'URL de connexion My Domain (par exemple *`MyDomainName--SandboxName.sandbox.my.salesforce.com`*) comme URL d'authentification.

C. Choisissez Enregistrer l'authentification.

d. Virtual Private Cloud (VPC) —Vous pouvez choisir d'utiliser un VPC. Dans ce cas, vous devez ajouter des sous-réseaux et des groupes de sécurité VPC.

e. Identity Crawler : spécifiez s'il faut activer l'explorateur Amazon Kendra d'identité. Le robot d'exploration d'identité utilise les informations de la liste de contrôle d'accès (ACL) de vos documents pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Si vous disposez d'une ACL pour vos documents et que vous choisissez de l'utiliser, vous pouvez également choisir d'activer le robot d'exploration Amazon Kendra d'identité pour configurer le filtrage des résultats de recherche par contexte utilisateur. Sinon, si le robot d'identification est désactivé, tous les documents peuvent être consultés publiquement. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'exploration d'identité est désactivé, vous pouvez également utiliser l'PutPrincipalMappingAPI pour télécharger les informations d'accès des utilisateurs et des groupes afin de filtrer le contexte utilisateur.

f. IAM rôle —Choisissez un IAM rôle existant ou créez-en un nouveau IAM pour accéder aux informations d'identification de votre référentiel et indexer le contenu.

> **Note**
>
> IAM les rôles utilisés pour les index ne peuvent pas être utilisés pour les sources de données. Si vous ne savez pas si un rôle existant est utilisé pour un index ou une FAQ, choisissez Créer un nouveau rôle pour éviter les erreurs.

g. Choisissez Suivant.

7. Sur la page Configurer les paramètres de synchronisation, entrez les informations suivantes :

   a. Pour analyser les pièces jointes : sélectionnez cette option pour analyser tous les objets Salesforce attachés.

   b. Pour les objets standard, les objets standard avec pièces jointes et les objets standard sans pièce jointe et articles de connaissances, sélectionnez les entités ou les types de contenu Salesforce que vous souhaitez analyser.

   c. Vous devez fournir des informations de configuration pour indexer au moins un objet standard, un article de connaissances ou un fil de discussion. Si vous choisissez d'explorer les articles de connaissances, vous devez spécifier les types d'articles de connaissances à indexer. Vous pouvez choisir les versions publiées, archivées, les brouillons et les pièces jointes.

      Filtre Regex : spécifiez un modèle d'expression régulière pour inclure des éléments de catalogue spécifiques.

8. Pour une configuration supplémentaire :

   - Informations ACL Toutes les listes de contrôle d'accès sont incluses par défaut. La désélection d'une liste de contrôle d'accès rendra publics tous les fichiers de cette catégorie.
   - Modèles Regex : ajoutez des modèles d'expressions régulières pour inclure ou exclure certains fichiers. Vous pouvez ajouter jusqu'à 100 motifs.

   Mode de synchronisation : choisissez la manière dont vous souhaitez mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation.

   - Synchronisation complète : Indexez tout le contenu fraîchement, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
   - Nouvelle synchronisation modifiée : Indexez uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

- Synchronisation nouvelle, modifiée ou supprimée : Indexez uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

9. Choisissez Suivant.
10. Sur la page Définir les mappages de champs, entrez les informations suivantes :

    a. Pour l'article de connaissance standard, les pièces jointes aux objets standard et les mappages de champs suggérés supplémentaires, sélectionnez parmi les champs de source de données par défaut Amazon Kendra générés que vous souhaitez mapper à votre index.

    > **Note**
    >
    > Un index mappé à `_document_body` est requis. Vous ne pouvez pas modifier le mappage entre le `Salesforce ID` champ et le Amazon Kendra `_document_id` champ. Vous pouvez associer n'importe quel champ Salesforce au titre du document ou aux champs d' reserved/default index Amazon Kendra du corps du document.
    > Si vous associez un champ Salesforce aux champs du titre et du corps du document Amazon Kendra, Amazon Kendra utilisera les données du titre et des champs du corps du document dans les réponses de recherche.

    b. Ajouter un champ —Pour ajouter des champs de source de données personnalisés afin de créer un nom de champ d'index à mapper et le type de données du champ.

    c. Choisissez Suivant.

11. Sur la page Réviser et créer, vérifiez que les informations saisies sont correctes, puis sélectionnez Ajouter une source de données. Vous pouvez également choisir de modifier vos informations depuis cette page. Votre source de données apparaîtra sur la page Sources de données une fois que la source de données aura été ajoutée avec succès.

API

Pour vous connecter Amazon Kendra à Salesforce

Vous devez spécifier un JSON du schéma de source de données à l'aide de l'TemplateConfigurationAPI. Vous devez fournir les informations suivantes :

- Source de données —Spécifiez le type de source de données tel que `SALESFORCEV2` lorsque vous utilisez le schéma TemplateConfigurationJSON. Spécifiez également la source de données `TEMPLATE` lorsque vous appelez l'CreateDataSourceAPI.
- URL de l'hôte : spécifiez l'URL de l'hôte de l'instance Salesforce.
- Mode de synchronisation : spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation. Vous pouvez choisir entre :
  - `FORCED_FULL_CRAWL`pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
  - `FULL_CRAWL`pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
  - `CHANGE_LOG`pour indexer uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
- Nom de ressource Amazon secret (ARN) : indiquez le nom de ressource Amazon (ARN) d'un Secrets Manager secret contenant les informations d'authentification de votre compte Salesforce. Le secret est stocké dans une structure JSON avec les clés suivantes :

```
{
    "authenticationUrl": "OAUTH endpoint that Amazon Kendra connects to get an
 OAUTH token",
    "consumerKey": "Application public key generated when you created your
 Salesforce application",
    "consumerSecret": "Application private key generated when you created your
 Salesforce application",
    "password": "Password associated with the user logging in to the Salesforce
 instance",
```

```
    "securityToken": "Token associated with the user account logging in to the
 Salesforce instance",
    "username": "User name of the user logging in to the Salesforce instance"
}
```

- IAM role —Spécifiez à quel `RoleArn` moment vous appelez `CreateDataSource` pour fournir à un IAM rôle les autorisations d'accéder à votre Secrets Manager secret et d'appeler le public requis APIs pour le connecteur Salesforce et Amazon Kendra. Pour plus d'informations, consultez la section IAM Rôles pour les sources de données Salesforce.

Vous pouvez également ajouter les fonctionnalités optionnelles suivantes :

- Virtual Private Cloud (VPC) **`VpcConfiguration`** —Spécifiez le moment de votre appel. `CreateDataSource` Pour de plus amples informations, veuillez consulter Configuration Amazon Kendra pour utiliser un Amazon VPC.
- Filtres d'inclusion et d'exclusion : vous pouvez spécifier si vous souhaitez inclure ou exclure certains documents, comptes, campagnes, cas, contacts, prospects, opportunités, solutions, tâches, groupes, chatteurs et fichiers d'entités personnalisés.

> **Note**
>
> La plupart des sources de données utilisent des modèles d'expressions régulières, qui sont des modèles d'inclusion ou d'exclusion appelés filtres. Si vous spécifiez un filtre d'inclusion, seul le contenu correspondant au filtre d'inclusion est indexé. Tout document qui ne correspond pas au filtre d'inclusion n'est pas indexé. Si vous spécifiez un filtre d'inclusion et d'exclusion, les documents correspondant au filtre d'exclusion ne sont pas indexés, même s'ils correspondent au filtre d'inclusion.

- Identity Crawler : spécifiez s'il faut activer l'explorateur Amazon Kendra d'identité. Le robot d'exploration d'identité utilise les informations de la liste de contrôle d'accès (ACL) de vos documents pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Si vous disposez d'une ACL pour vos documents et que vous choisissez de l'utiliser, vous pouvez également choisir d'activer le robot d'exploration Amazon Kendra d'identité pour configurer le filtrage des résultats de recherche par contexte utilisateur. Sinon, si le robot d'identification est désactivé, tous les documents peuvent être consultés publiquement. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'exploration d'identité est désactivé, vous pouvez également utiliser l'PutPrincipalMappingAPI

pour télécharger les informations d'accès des utilisateurs et des groupes afin de filtrer le contexte utilisateur.

- Mappages de champs : choisissez de mapper les champs de votre source de données Salesforce à vos champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.

> **Note**
>
> Le champ du corps du document ou l'équivalent du corps du document pour vos documents est requis pour Amazon Kendra effectuer une recherche dans vos documents. Vous devez associer le nom du champ du corps du document dans votre source de données au nom du champ d'index`_document_body`. Tous les autres champs sont facultatifs.

> **Note**
>
> Un index mappé à `_document_body` est requis. Vous ne pouvez pas modifier le mappage entre le `Salesforce ID` champ et le Amazon Kendra `_document_id` champ. Vous pouvez associer n'importe quel champ Salesforce au titre du document ou aux champs d' reserved/default index Amazon Kendra du corps du document.
> Si vous associez un champ Salesforce aux champs du titre et du corps du document Amazon Kendra, Amazon Kendra utilisera les données du titre et des champs du corps du document dans les réponses de recherche.

Pour obtenir la liste des autres clés JSON importantes à configurer, consultez le schéma du modèle Salesforce.

En savoir plus

Pour en savoir plus sur l'intégration Amazon Kendra à votre source de données Salesforce, consultez :

- Annonce de la mise à jour du connecteur Salesforce (V2) pour Amazon Kendra

Remarques

- Lorsque les listes de contrôle d'accès (ACLs) sont activées, l'option « Synchroniser uniquement le contenu nouveau ou modifié » n'est pas disponible en raison des limites de l'API Salesforce. Nous vous recommandons d'utiliser plutôt les modes « Synchronisation complète » ou « Synchronisation du contenu nouveau, modifié ou supprimé », ou de les désactiver ACLs si vous devez utiliser ce mode de synchronisation.

ServiceNow

ServiceNow fournit un système de gestion des services basé sur le cloud pour créer et gérer des flux de travail au niveau de l'organisation, tels que les services informatiques, les systèmes de billetterie et le support. Vous pouvez l'utiliser Amazon Kendra pour indexer vos ServiceNow catalogues, articles de connaissances, incidents et leurs pièces jointes.

Vous pouvez vous connecter Amazon Kendra à votre source de ServiceNow données à l'aide de la Amazon Kendra console, de l'TemplateConfigurationAPI ou de l'ServiceNowConfigurationAPI.

Amazon Kendra possède deux versions du ServiceNow connecteur. Les fonctionnalités prises en charge par chaque version incluent :

ServiceNow connecteur V1.0/API ServiceNowConfiguration

- Mappages de champs
- ServiceNow versions d'instance : Londres, Autres
- Filtres d'inclusion/exclusion

ServiceNow connecteur V2.0/API TemplateConfiguration

- Mappages de champs
- Contrôle d'accès utilisateur
- Filtres d'inclusion/exclusion
- Synchronisation complète et incrémentielle du contenu
- ServiceNow versions d'instance : Rome, Sandiego, Tokyo, autres
- Cloud privé virtuel (VPC)

> **Note**
>
> ServiceNow le connecteur V1.0/ ServiceNowConfiguration API a pris fin en 2023. Nous vous recommandons de migrer vers ou d'utiliser le ServiceNow connecteur TemplateConfiguration V2.0/API.

Pour résoudre les problèmes liés à votre connecteur de source de Amazon Kendra ServiceNow données, consultezDépannage des sources de données.

Rubriques

- ServiceNow connecteur V1.0
- ServiceNow connecteur V2.0
- Spécification des documents à indexer à l'aide d'une requête

ServiceNow connecteur V1.0

ServiceNow fournit un système de gestion des services basé sur le cloud pour créer et gérer des flux de travail au niveau de l'organisation, tels que les services informatiques, les systèmes de billetterie et le support. Vous pouvez l'utiliser Amazon Kendra pour indexer vos ServiceNow catalogues, vos articles de connaissances et leurs pièces jointes.

> **Note**
>
> ServiceNow le connecteur V1.0/ ServiceNowConfiguration API a pris fin en 2023. Nous vous recommandons de migrer vers ou d'utiliser le ServiceNow connecteur TemplateConfiguration V2.0/API.

Pour résoudre les problèmes liés à votre connecteur de source de Amazon Kendra ServiceNow données, consultezDépannage des sources de données.

Rubriques

- Fonctionnalités prises en charge
- Prérequis
- Instructions de connexion

- En savoir plus

Fonctionnalités prises en charge

Amazon Kendra ServiceNow le connecteur de source de données prend en charge les fonctionnalités suivantes :

- ServiceNow versions d'instance : Londres, Autres
- Modèles d'inclusion/exclusion : catalogues de services, articles de connaissances et leurs pièces jointes

Prérequis

Avant de pouvoir Amazon Kendra indexer votre source de ServiceNow données, apportez ces modifications à vos AWS comptes ServiceNow and.

Dans ServiceNow, assurez-vous d'avoir :

- J'ai créé un compte ServiceNow administrateur et j'ai créé une ServiceNow instance.
- Vous avez copié l'URL de l'hôte de votre ServiceNow instance. Par exemple, si l'URL de l'instance est*https://your-domain.service-now.com*, le format de l'URL hôte que vous entrez est le suivant*your-domain.service-now.com*.
- Vous avez noté vos informations d'authentification de base contenant un nom d'utilisateur et un mot de passe pour vous Amazon Kendra permettre de vous connecter à votre ServiceNow instance.

  > **Note**
  >
  > Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

- Facultatif : configuration d'un jeton d'identification OAuth 2.0 capable d'identifier Amazon Kendra et de générer un nom d'utilisateur, un mot de passe, un identifiant client et un secret client. Le nom d'utilisateur et le mot de passe doivent permettre d'accéder à la base de ServiceNow

connaissances et au catalogue de services. Consultez ServiceNow la documentation sur l'authentification OAuth 2.0 pour plus d'informations.

- Les autorisations suivantes ont été ajoutées :
  - kb_category
  - kb_knowledge
  - kb_knowledge_base
  - kb_uc_ne peut pas lire le mtom
  - kb_uc_can_read_mtom
  - sc_catalog
  - sc_category
  - sc_cat_item
  - sys_attachment
  - sys_attachment_doc
  - rôle_utilisateur système
- Il est vérifié que chaque document est unique dans ServiceNow et entre les autres sources de données que vous prévoyez d'utiliser pour le même index. Chaque source de données que vous souhaitez utiliser pour un index ne doit pas contenir le même document dans toutes les sources de données. IDs Les documents sont globaux par rapport à un index et doivent être uniques par index.

Dans votre Compte AWS, assurez-vous d'avoir :

- Création d'un Amazon Kendra index et, si vous utilisez l'API, notez l'ID de l'index.
- Vous avez créé un IAM rôle pour votre source de données et, si vous utilisez l'API, notez l'ARN du IAM rôle.

> **Note**
>
> Si vous modifiez votre type d'authentification et vos informations d'identification, vous devez mettre à jour votre IAM rôle pour accéder au bon identifiant AWS Secrets Manager secret.

- Stockez vos informations d' ServiceNow authentification dans un AWS Secrets Manager secret et, si vous utilisez l'API, notez l'ARN du secret.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

Si vous n'avez pas de IAM rôle ou de secret existant, vous pouvez utiliser la console pour créer un nouveau IAM rôle et un nouveau Secrets Manager secret lorsque vous connectez votre source de ServiceNow données à Amazon Kendra. Si vous utilisez l'API, vous devez fournir l'ARN d'un IAM rôle et d'un Secrets Manager secret existants, ainsi qu'un identifiant d'index.

Instructions de connexion

Pour vous connecter Amazon Kendra à votre source de ServiceNow données, vous devez fournir les informations nécessaires sur votre source de ServiceNow données afin de Amazon Kendra pouvoir accéder à vos données. Si vous n'avez pas encore configuré ServiceNow pour Amazon Kendra voirPrérequis.

Console

Pour vous connecter Amazon Kendra à ServiceNow

1. Connectez-vous à la console AWS de gestion et Amazon Kendra ouvrez-la.
2. Dans le volet de navigation de gauche, choisissez Index, puis choisissez l'index que vous souhaitez utiliser dans la liste des index.

   > **Note**
   >
   > Vous pouvez choisir de configurer ou de modifier vos paramètres de contrôle d'accès utilisateur dans les paramètres de l'index.

3. Sur la page de démarrage, choisissez Ajouter une source de données.
4. Sur la page Ajouter une source de données, choisissez le ServiceNowconnecteur V1.0, puis sélectionnez Ajouter une source de données.
5. Sur la page Spécifier les détails de la source de données, entrez les informations suivantes :

a. Dans Nom et description, pour Nom de la source de données : entrez le nom de votre source de données. Vous pouvez inclure des traits d'union, mais pas des espaces.

b. (Facultatif) Description : entrez une description facultative pour votre source de données.

c. Dans la langue par défaut : choisissez une langue pour filtrer vos documents pour l'index. Sauf indication contraire, la langue par défaut est l'anglais. La langue spécifiée dans les métadonnées du document remplace la langue sélectionnée.

d. Dans Balises, pour Ajouter une nouvelle balise : incluez des balises facultatives pour rechercher et filtrer vos ressources ou suivre vos AWS coûts.

e. Choisissez Suivant.

6. Sur la page Définir l'accès et la sécurité, entrez les informations suivantes :

a. ServiceNow host —Entrez l'URL de l' ServiceNowhôte.

b. ServiceNow version —Sélectionnez votre ServiceNow version.

c. Choisissez entre l'authentification de base et l'authentification Oauth 2.0 en fonction de votre cas d'utilisation.

d. AWS Secrets Manager secret —Choisissez un secret existant ou créez-en un nouveau Secrets Manager pour stocker vos informations ServiceNow d'authentification. Si vous choisissez de créer un nouveau secret, une fenêtre AWS Secrets Manager secrète s'ouvre.

i. Nom secret : le nom de votre secret. Le préfixe « AmazonKendra - ServiceNow - » est automatiquement ajouté à votre nom secret.

ii. Si vous utilisez l'authentification de base, entrez le nom secret, le nom d'utilisateur et le mot de passe de votre compte. ServiceNow

Si vous utilisez l' OAuth2 authentification, entrez le nom secret, le nom d'utilisateur, le mot de passe, l'identifiant client et le secret client que vous avez créés dans votre compte. ServiceNow

iii. Choisissez Enregistrer et ajoutez un secret.

e. IAM rôle —Choisissez un IAM rôle existant ou créez-en un nouveau IAM pour accéder aux informations d'identification de votre référentiel et indexer le contenu.

> **Note**
>
> IAM les rôles utilisés pour les index ne peuvent pas être utilisés pour les sources de données. Si vous ne savez pas si un rôle existant est utilisé pour un index ou une FAQ, choisissez Créer un nouveau rôle pour éviter les erreurs.

f. Choisissez Suivant.

7. Sur la page Configurer les paramètres de synchronisation, entrez les informations suivantes :

   a. Inclure les articles de connaissances : choisissez d'indexer les articles de connaissances.

   b. Type d'articles de connaissances : choisissez entre Inclure uniquement les articles publics et Inclure les articles en fonction d'une requête de ServiceNow filtrage basée sur votre cas d'utilisation. Si vous sélectionnez Inclure les articles en fonction d'une requête de ServiceNow filtre, vous devez saisir une requête de filtre copiée depuis votre ServiceNow compte.

   c. Inclure les pièces jointes aux articles de connaissances : choisissez d'indexer les pièces jointes aux articles de connaissances. Vous pouvez également sélectionner des types de fichiers spécifiques à indexer.

   d. Inclure les éléments du catalogue : choisissez d'indexer les éléments du catalogue.

   e. Inclure les pièces jointes aux éléments du catalogue : choisissez d'indexer les pièces jointes aux éléments du catalogue. Vous pouvez également sélectionner des types de fichiers spécifiques à indexer.

   f. Fréquence : fréquence Amazon Kendra de synchronisation avec votre source de données.

   g. Choisissez Suivant.

8. Sur la page Définir les mappages de champs, entrez les informations suivantes :

   a. Articles de connaissances et catalogue de services : sélectionnez parmi les champs de source de données par défaut Amazon Kendra générés et les mappages de champs suggérés supplémentaires que vous souhaitez mapper à votre index.

   b. Ajouter un champ —Pour ajouter des champs de source de données personnalisés afin de créer un nom de champ d'index à mapper et le type de données du champ.

   c. Choisissez Suivant.

9. Sur la page Réviser et créer, vérifiez que les informations saisies sont correctes, puis sélectionnez Ajouter une source de données. Vous pouvez également choisir de modifier vos informations depuis cette page. Votre source de données apparaîtra sur la page Sources de données une fois que la source de données aura été ajoutée avec succès.

API

Pour vous connecter Amazon Kendra à ServiceNow

Vous devez spécifier les éléments suivants à l'aide de l'ServiceNowConfiguration API :

- URL de la source de données : spécifiez l' ServiceNow URL. Le point de terminaison de l'hôte doit ressembler à ce qui suit :*your-domain.service-now.com*.
- Instance hôte de source de données : spécifiez la version de l'instance ServiceNow hôte sous la forme `LONDON` ou`OTHERS`.
- Nom de ressource Amazon (ARN) secret : indiquez le nom de ressource Amazon (ARN) d'un Secrets Manager secret contenant les informations d'authentification que vous avez créées dans votre ServiceNow compte.

  Si vous utilisez l'authentification de base, le secret est stocké dans une structure JSON avec les clés suivantes :

  ```
  {
      "username": "user name",
      "password": "password"
  }
  ```

  Si vous utilisez l' OAuth2 authentification, le secret est stocké dans une structure JSON avec les clés suivantes :

  ```
  {
      "username": "user name",
      "password": "password",
      "clientId": "client id",
      "clientSecret": "client secret"
  }
  ```

- IAM role —Spécifiez le `RoleArn` moment où vous appelez `CreateDataSource` pour fournir à un IAM rôle les autorisations d'accéder à votre Secrets Manager secret et d'appeler le public

requis APIs pour le ServiceNow connecteur et Amazon Kendra. Pour plus d'informations, consultez la section IAM Rôles des sources de ServiceNow données.

Vous pouvez également ajouter les fonctionnalités optionnelles suivantes :

- Mappages de champs : choisissez de mapper les champs de votre source de ServiceNow données à vos champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.

  > **Note**
  >
  > Le champ du corps du document ou l'équivalent du corps du document pour vos documents est requis pour Amazon Kendra effectuer une recherche dans vos documents. Vous devez associer le nom du champ du corps du document dans votre source de données au nom du champ d'index`_document_body`. Tous les autres champs sont facultatifs.

- Filtres d'inclusion et d'exclusion : indiquez si vous souhaitez inclure ou exclure certaines pièces jointes de catalogues et d'articles de connaissances.

  > **Note**
  >
  > La plupart des sources de données utilisent des modèles d'expressions régulières, qui sont des modèles d'inclusion ou d'exclusion appelés filtres. Si vous spécifiez un filtre d'inclusion, seul le contenu correspondant au filtre d'inclusion est indexé. Tout document qui ne correspond pas au filtre d'inclusion n'est pas indexé. Si vous spécifiez un filtre d'inclusion et d'exclusion, les documents correspondant au filtre d'exclusion ne sont pas indexés, même s'ils correspondent au filtre d'inclusion.

- Paramètres d'indexation : vous pouvez également choisir de spécifier si vous souhaitez :
  - Indexez les articles de connaissances et les catalogues de services, ou les deux. Si vous choisissez d'indexer des articles de connaissances et des éléments du catalogue de services, vous devez fournir le nom du ServiceNow champ mappé au champ de contenu du document d'index dans l' Amazon Kendra index.
  - Indexez les pièces jointes aux articles de connaissances et aux éléments du catalogue.
  - Utilisez une ServiceNow requête qui sélectionne des documents dans une ou plusieurs bases de connaissances. Les bases de connaissances peuvent être publiques ou privées. Pour plus

d'informations, veuillez consulter la rubrique Spécification des documents à indexer avec une requête.

En savoir plus

Pour en savoir plus sur l'intégration Amazon Kendra à votre source de ServiceNow données, consultez :

- Commencer à utiliser le connecteur Amazon Kendra ServiceNow en ligne

ServiceNow connecteur V2.0

ServiceNow fournit un système de gestion des services basé sur le cloud pour créer et gérer des flux de travail au niveau de l'organisation, tels que les services informatiques, les systèmes de billetterie et le support. Vous pouvez l'utiliser Amazon Kendra pour indexer vos ServiceNow catalogues, articles de connaissances, incidents et leurs pièces jointes.

Pour résoudre les problèmes liés à votre connecteur de source de Amazon Kendra ServiceNow données, consultezDépannage des sources de données.

Rubriques

- Fonctionnalités prises en charge
- Prérequis
- Instructions de connexion
- En savoir plus

Fonctionnalités prises en charge

Amazon Kendra ServiceNow le connecteur de source de données prend en charge les fonctionnalités suivantes :

- Mappages de champs
- Contrôle d'accès utilisateur
- Filtres d'inclusion/exclusion
- Synchronisation complète et incrémentielle du contenu
- ServiceNow versions d'instance : Rome, Sandiego, Tokyo, autres

- Cloud privé virtuel (VPC)

Prérequis

Avant de pouvoir Amazon Kendra indexer votre source de ServiceNow données, apportez ces modifications à vos AWS comptes ServiceNow and.

Dans ServiceNow, assurez-vous d'avoir :

- Vous avez créé une instance de développeur personnelle ou d'entreprise et disposez d'une ServiceNow instance dotée d'un rôle administratif.
- Vous avez copié l'URL de l'hôte de votre ServiceNow instance. Le format de l'URL d'hôte que vous entrez est*your-domain.service-now.com*. Vous avez besoin de l'URL de votre ServiceNow instance pour vous connecter Amazon Kendra.
- Vous avez pris note de vos informations d'authentification de base, à savoir un nom d'utilisateur et un mot de passe Amazon Kendra pour autoriser la connexion à votre ServiceNow instance.

  > **Note**
  >
  > Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

- Facultatif : informations d'identification client OAuth 2.0 configurées qui peuvent être identifiées à l' Amazon Kendra aide d'un nom d'utilisateur, d'un mot de passe, d'un identifiant client généré et d'un secret client. Consultez ServiceNow la documentation sur l'authentification OAuth 2.0 pour plus d'informations.
- Les autorisations suivantes ont été ajoutées :
  - kb_category
  - kb_knowledge
  - kb_knowledge_base
  - kb_uc_ne peut pas lire le mtom
  - kb_uc_can_read_mtom
  - sc_catalog

  - sc_category
  - sc_cat_item
  - sys_attachment
  - sys_attachment_doc
  - rôle_utilisateur système
- Il est vérifié que chaque document est unique dans ServiceNow et entre les autres sources de données que vous prévoyez d'utiliser pour le même index. Chaque source de données que vous souhaitez utiliser pour un index ne doit pas contenir le même document dans toutes les sources de données. IDs Les documents sont globaux par rapport à un index et doivent être uniques par index.

Dans votre Compte AWS, assurez-vous d'avoir :

- Création d'un Amazon Kendra index et, si vous utilisez l'API, notez l'ID de l'index.
- Vous avez créé un IAM rôle pour votre source de données et, si vous utilisez l'API, notez l'ARN du IAM rôle.

> **Note**
>
> Si vous modifiez votre type d'authentification et vos informations d'identification, vous devez mettre à jour votre IAM rôle pour accéder au bon identifiant AWS Secrets Manager secret.

- Stockez vos informations d' ServiceNow authentification dans un AWS Secrets Manager secret et, si vous utilisez l'API, notez l'ARN du secret.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

Si vous n'avez pas de IAM rôle ou de secret existant, vous pouvez utiliser la console pour créer un nouveau IAM rôle et un nouveau Secrets Manager secret lorsque vous connectez votre source de

ServiceNow données à Amazon Kendra. Si vous utilisez l'API, vous devez fournir l'ARN d'un IAM rôle et d'un Secrets Manager secret existants, ainsi qu'un identifiant d'index.

Instructions de connexion

Pour vous connecter Amazon Kendra à votre source de ServiceNow données, vous devez fournir les informations nécessaires sur votre source de ServiceNow données afin de Amazon Kendra pouvoir accéder à vos données. Si vous n'avez pas encore configuré ServiceNow pour Amazon Kendra voirPrérequis.

Console

Pour vous connecter Amazon Kendra à ServiceNow

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la.
2. Dans le volet de navigation de gauche, choisissez Index, puis choisissez l'index que vous souhaitez utiliser dans la liste des index.

   > **Note**
   >
   > Vous pouvez choisir de configurer ou de modifier vos paramètres de contrôle d'accès utilisateur dans les paramètres de l'index.

3. Sur la page de démarrage, choisissez Ajouter une source de données.
4. Sur la page Ajouter une source de données, choisissez ServiceNow connecteur, puis sélectionnez Ajouter un connecteur. Si vous utilisez la version 2 (le cas échéant), choisissez le ServiceNow connecteur avec la balise « V2.0 ».
5. Sur la page Spécifier les détails de la source de données, entrez les informations suivantes :

   a. Dans Nom et description, pour Nom de la source de données : entrez le nom de votre source de données. Vous pouvez inclure des traits d'union, mais pas des espaces.
   b. (Facultatif) Description : entrez une description facultative pour votre source de données.
   c. Dans la langue par défaut : choisissez une langue pour filtrer vos documents pour l'index. Sauf indication contraire, la langue par défaut est l'anglais. La langue spécifiée dans les métadonnées du document remplace la langue sélectionnée.
   d. Dans Balises, pour Ajouter une nouvelle balise : incluez des balises facultatives pour rechercher et filtrer vos ressources ou suivre vos AWS coûts.
   e. Choisissez Suivant.

6. Sur la page Définir l'accès et la sécurité, entrez les informations suivantes :

   a. ServiceNow host —Entrez l'URL de l' ServiceNowhôte. Le format de l'URL d'hôte que vous entrez est*your-domain.service-now.com*.

   b. ServiceNow version —Sélectionnez la version de votre ServiceNow instance. Vous pouvez choisir entre Rome, Sandiego, Tokyo ou autres.

   c. Autorisation : activez ou désactivez les informations de la liste de contrôle d'accès (ACL) pour vos documents, si vous disposez d'une ACL et que vous souhaitez l'utiliser pour le contrôle d'accès. L'ACL indique les documents auxquels les utilisateurs et les groupes peuvent accéder. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.

   d. Authentification : choisissez entre l'authentification de base et l'authentification Oauth 2.0.

   e. AWS Secrets Manager secret —Choisissez un secret existant ou créez-en un nouveau Secrets Manager pour stocker vos informations ServiceNow d'authentification. Si vous choisissez de créer un nouveau secret, une fenêtre AWS Secrets Manager secrète s'ouvre. Entrez les informations suivantes dans la fenêtre :

      i. Nom secret : le nom de votre secret. Le préfixe « AmazonKendra - ServiceNow - » est automatiquement ajouté à votre nom secret.

      ii. Si vous utilisez l'authentification de base, entrez le nom secret, le nom d'utilisateur et le mot de passe de votre compte. ServiceNow

      Si vous utilisez l'authentification OAuth2 .0, entrez le nom secret, le nom d'utilisateur, le mot de passe, l'identifiant client et le secret client que vous avez créés dans votre compte. ServiceNow

      iii. Enregistrez et ajoutez votre secret.

   f. Virtual Private Cloud (VPC) —Vous pouvez choisir d'utiliser un VPC. Dans ce cas, vous devez ajouter des sous-réseaux et des groupes de sécurité VPC.

   g. Identity Crawler : spécifiez s'il faut activer l'explorateur Amazon Kendra d'identité. Le robot d'exploration d'identité utilise les informations de la liste de contrôle d'accès (ACL) de vos documents pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Si vous disposez d'une ACL pour vos documents et que vous choisissez de l'utiliser, vous pouvez également choisir d'activer le robot d'exploration Amazon Kendra d'identité pour configurer le filtrage des résultats

de recherche par contexte utilisateur. Sinon, si le robot d'identification est désactivé, tous les documents peuvent être consultés publiquement. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'exploration d'identité est désactivé, vous pouvez également utiliser l'PutPrincipalMappingAPI pour télécharger les informations d'accès des utilisateurs et des groupes afin de filtrer le contexte utilisateur.

h. IAM rôle —Choisissez un IAM rôle existant ou créez-en un nouveau IAM pour accéder aux informations d'identification de votre référentiel et indexer le contenu.

> **Note**
>
> IAM les rôles utilisés pour les index ne peuvent pas être utilisés pour les sources de données. Si vous ne savez pas si un rôle existant est utilisé pour un index ou une FAQ, choisissez Créer un nouveau rôle pour éviter les erreurs.

i. Choisissez Suivant.

7. Sur la page Configurer les paramètres de synchronisation, entrez les informations suivantes :

a. Pour les articles de Knowledge, choisissez l'une des options suivantes :

- Articles de connaissances : choisissez d'indexer les articles de connaissances.
- Pièces jointes aux articles de connaissances : choisissez d'indexer les pièces jointes aux articles de connaissances.
- Type d'articles de connaissances : choisissez entre uniquement les articles publics et les articles de connaissances en fonction d'une requête de ServiceNow filtrage basée sur votre cas d'utilisation. Si vous sélectionnez Inclure les articles en fonction d'une requête de ServiceNow filtre, vous devez saisir une requête de filtre copiée depuis votre ServiceNow compte. Les exemples de requêtes de filtrage incluent :*workflow_state=draft^EQ*,*kb_knowledge_base=dfc19531bf2021003f07e ISNOTEMPTY^EQ*,*article_type=text^active=true^EQ*.

> **Important**
>
> Si vous choisissez d'explorer uniquement les articles publics, n' Amazon Kendra explore que les articles de connaissances auxquels un rôle d'accès public a été attribué dans. ServiceNow

- Inclure des articles en fonction d'un filtre de brève description : spécifiez des modèles d'expressions régulières pour inclure ou exclure des articles spécifiques.

b. Pour les articles du catalogue de services :

- Articles du catalogue de services : choisissez d'indexer les articles du catalogue de services.
- Pièces jointes des articles du catalogue de services : choisissez d'indexer les pièces jointes des articles du catalogue de services.
- Éléments du catalogue de services actifs : choisissez d'indexer les éléments du catalogue de services actifs.
- Éléments du catalogue de services inactifs : choisissez d'indexer les éléments du catalogue de services inactifs.
- Requête de filtre : choisissez d'inclure les éléments du catalogue de services en fonction d'un filtre défini dans votre ServiceNow instance. Les exemples de requêtes de filtrage incluent :*short_descriptionLIKEAccess^category=2809952237b1300054b6a3549c*
- Inclure les éléments du catalogue de services en fonction d'un filtre de brève description : spécifiez un modèle d'expression régulière pour inclure des articles de catalogue spécifiques.

c. Pour les incidents :

- Incidents : choisissez d'indexer les incidents de service.
- Pièces jointes aux incidents : choisissez d'indexer les pièces jointes aux incidents.
- Incidents actifs : choisissez d'indexer les incidents actifs.
- Incidents inactifs : choisissez d'indexer les incidents inactifs.
- Type d'incident actif : choisissez entre Tous les incidents, Incidents ouverts, Incidents non assignés et Incidents résolus en fonction de votre cas d'utilisation.
- Requête de filtre : choisissez d'inclure les incidents en fonction d'un filtre défini dans votre ServiceNow instance. Les exemples de requêtes de filtrage incluent :*short_descriptionLIKETest^urgency=3^state=1^EQ,priority=2^categ* .
- Inclure les incidents en fonction d'un filtre de brève description : spécifiez un modèle d'expression régulière pour inclure des incidents spécifiques.

d. Pour une configuration supplémentaire :

- Informations ACL : les listes de contrôle d'accès pour les entités que vous avez sélectionnées sont incluses par défaut. La désélection d'une liste de contrôle d'accès rendra publics tous les fichiers de cette catégorie. Les options ACL sont automatiquement désactivées pour les entités non sélectionnées. Pour les articles publics, l'ACL n'est pas appliquée.
- Pour la taille de fichier maximale : spécifiez la taille limite de fichier MBs qu'Amazon Kendra explorera. Amazon Kendra explorera uniquement les fichiers dans la limite de taille que vous avez définie. La taille de fichier par défaut est de 50 Mo. La taille maximale du fichier doit être supérieure à 0 Mo et inférieure ou égale à 50 Mo.
- Modèles d'expressions régulières pour pièces jointes : ajoutez des modèles d'expressions régulières pour inclure ou exclure certains fichiers joints de catalogues, d'articles de connaissances et d'incidents. Vous pouvez ajouter jusqu'à 100 motifs.

e. Mode de synchronisation : choisissez la manière dont vous souhaitez mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation.

- Synchronisation complète : Indexez tout le contenu fraîchement, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
- Synchronisation nouvelle, modifiée ou supprimée : Indexez uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

f. Dans Calendrier d'exécution de la synchronisation, pour Fréquence : choisissez la fréquence à laquelle vous souhaitez synchroniser le contenu de votre source de données et mettre à jour votre index.

g. Choisissez Suivant.

8. Sur la page Définir les mappages de champs, entrez les informations suivantes :

a. Mappages de champs par défaut : sélectionnez parmi les champs de source de données par défaut Amazon Kendra générés que vous souhaitez mapper à votre index.

b. Ajouter un champ —Pour ajouter des champs de source de données personnalisés afin de créer un nom de champ d'index à mapper et le type de données du champ.

c. Choisissez Suivant.

9. Sur la page Réviser et créer, vérifiez que les informations saisies sont correctes, puis sélectionnez Ajouter une source de données. Vous pouvez également choisir de modifier vos informations depuis cette page. Votre source de données apparaîtra sur la page Sources de données une fois que la source de données aura été ajoutée avec succès.

API

Pour vous connecter Amazon Kendra à ServiceNow

Vous devez spécifier un JSON du schéma de source de données à l'aide de l'TemplateConfigurationAPI. Vous devez fournir les informations suivantes :

- Source de données —Spécifiez le type de source de données comme `SERVICENOWV2` lorsque vous utilisez TemplateConfigurationSchéma JSON. Spécifiez également la source de données comme `TEMPLATE` lorsque vous appelez le CreateDataSourceAPI.
- URL de l'hôte : spécifiez la version de l'instance ServiceNow hôte. Par exemple, *`your-domain.service-now.com`*.
- Type d'authentification : spécifiez le type d'authentification que vous utilisez, que ce soit `basicAuth` ou `OAuth2` pour votre ServiceNow instance.
- ServiceNow version de l'instance —Spécifiez l' ServiceNow instance que vous utilisez`Tokyo`, si`Sandiego`,`Rome`, ou`Others`.
- Mode de synchronisation : spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation. Vous pouvez choisir entre :
  - `FORCED_FULL_CRAWL`pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
  - `FULL_CRAWL`pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

- Nom de ressource Amazon (ARN) secret : indiquez le nom de ressource Amazon (ARN) d'un Secrets Manager secret contenant les informations d'authentification que vous avez créées dans votre ServiceNow compte.

  Si vous utilisez l'authentification de base, le secret est stocké dans une structure JSON avec les clés suivantes :

  ```
  {
      "username": "user name",
      "password": "password"
  }
  ```

- Si vous utilisez les informations d'identification du OAuth2 client, le secret est stocké dans une structure JSON avec les clés suivantes :

  ```
  {
      "username": "user name",
      "password": "password",
      "clientId": "client id",
      "clientSecret": "client secret"
  }
  ```

- IAM role —Spécifiez le `RoleArn` moment où vous appelez `CreateDataSource` pour fournir à un IAM rôle les autorisations d'accéder à votre Secrets Manager secret et d'appeler le public requis APIs pour le ServiceNow connecteur et Amazon Kendra. Pour plus d'informations, consultez la section IAM Rôles des sources de ServiceNow données.

Vous pouvez également ajouter les fonctionnalités optionnelles suivantes :

- Virtual Private Cloud (VPC) **`VpcConfiguration`** —Spécifiez le moment de votre appel. `CreateDataSource` Pour de plus amples informations, veuillez consulter Configuration Amazon Kendra pour utiliser un Amazon VPC.
- Filtres d'inclusion et d'exclusion : vous pouvez spécifier si vous souhaitez inclure ou exclure certains fichiers joints en utilisant les noms de fichiers et les types de fichiers des articles de connaissances, des catalogues de services et des incidents.

> **Note**
>
> La plupart des sources de données utilisent des modèles d'expressions régulières, qui sont des modèles d'inclusion ou d'exclusion appelés filtres. Si vous spécifiez un filtre d'inclusion, seul le contenu correspondant au filtre d'inclusion est indexé. Tout document qui ne correspond pas au filtre d'inclusion n'est pas indexé. Si vous spécifiez un filtre d'inclusion et d'exclusion, les documents correspondant au filtre d'exclusion ne sont pas indexés, même s'ils correspondent au filtre d'inclusion.

- Documents spécifiques à indexer : vous pouvez utiliser une ServiceNow requête pour spécifier les documents que vous souhaitez obtenir à partir d'une ou de plusieurs bases de connaissances, y compris les bases de connaissances privées. L'accès aux bases de connaissances est déterminé par l'utilisateur que vous utilisez pour vous connecter à l' ServiceNow instance. Pour plus d'informations, veuillez consulter la rubrique Spécification des documents à indexer avec une requête.
- Paramètres d'indexation : vous pouvez également choisir de spécifier si vous souhaitez :
  - Indexez les articles de connaissances, les catalogues de services et les incidents, ou tous ces éléments. Si vous choisissez d'indexer des articles de connaissances, des éléments du catalogue de services et des incidents, vous devez fournir le nom du ServiceNow champ mappé au champ de contenu du document d'index dans l' Amazon Kendra index.
  - Indexez les pièces jointes aux articles de connaissances, aux éléments du catalogue de services et aux incidents.
  - Incluez des articles de connaissances, des éléments du catalogue de services et des incidents en fonction du modèle de `short description` filtre.
  - Choisissez de filtrer les articles et incidents du catalogue de services actifs et inactifs.
  - Choisissez de filtrer les incidents en fonction du type d'incident.
  - Choisissez les entités dont l'ACL doit être analysée.
  - Vous pouvez utiliser une ServiceNow requête pour spécifier les documents que vous souhaitez obtenir à partir d'une ou de plusieurs bases de connaissances, y compris des bases de connaissances privées. L'accès aux bases de connaissances est déterminé par l'utilisateur que vous utilisez pour vous connecter à l' ServiceNowinstance. Pour plus d'informations, veuillez consulter la rubrique Spécification des documents à indexer avec une requête.

- Identity Crawler : spécifiez s'il faut activer l'explorateur Amazon Kendra d'identité. Le robot d'exploration d'identité utilise les informations de la liste de contrôle d'accès (ACL) de vos documents pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Si vous disposez d'une ACL pour vos documents et que vous choisissez de l'utiliser, vous pouvez également choisir d'activer le robot d'exploration Amazon Kendra d'identité pour configurer le filtrage des résultats de recherche par contexte utilisateur. Sinon, si le robot d'identification est désactivé, tous les documents peuvent être consultés publiquement. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'exploration d'identité est désactivé, vous pouvez également utiliser l'PutPrincipalMappingAPI pour télécharger les informations d'accès des utilisateurs et des groupes afin de filtrer le contexte utilisateur.
- Mappages de champs : choisissez de mapper les champs de votre source de ServiceNow données à vos champs d' Amazon Kendra index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.

> **Note**
>
> Le champ du corps du document ou l'équivalent du corps du document pour vos documents est requis pour Amazon Kendra effectuer une recherche dans vos documents. Vous devez associer le nom du champ du corps du document dans votre source de données au nom du champ d'index`_document_body`. Tous les autres champs sont facultatifs.

Pour obtenir la liste des autres clés JSON importantes à configurer, consultez le schéma du ServiceNow modèle.

En savoir plus

Pour en savoir plus sur l'intégration Amazon Kendra à votre source de ServiceNow données, consultez :

- Commencer à Amazon Kendra annoncer le ServiceNow connecteur mis à jour (V2) pour Amazon Kendra

Spécification des documents à indexer à l'aide d'une requête

Vous pouvez utiliser une ServiceNow requête pour spécifier les documents que vous souhaitez inclure dans un Amazon Kendra index. Lorsque vous utilisez une requête, vous pouvez spécifier plusieurs bases de connaissances, y compris des bases de connaissances privées. L'accès aux bases de connaissances est déterminé par l'utilisateur que vous utilisez pour vous connecter à l' ServiceNow instance.

Pour créer une requête, vous devez utiliser le générateur de ServiceNow requêtes. Vous pouvez utiliser le générateur pour créer la requête et vérifier qu'elle renvoie la bonne liste de documents.

Pour créer une requête à l'aide de la ServiceNow console

1. Connectez-vous à la ServiceNow console.
2. Dans le menu de gauche, choisissez Connaissances, puis Articles, puis choisissez Tout.
3. En haut de la page, choisissez l'icône du filtre.
4. Utilisez le générateur de requêtes pour créer la requête.
5. Lorsque la requête est terminée, cliquez avec le bouton droit sur la requête et choisissez Copier la requête pour copier la requête depuis le générateur de requêtes. Enregistrez cette requête pour l'utiliser dans Amazon Kendra.

Assurez-vous de ne modifier aucun paramètre de requête lorsque vous copiez la requête. Si l'un des paramètres de requête n'est pas reconnu, ServiceNow traite le paramètre comme vide et ne l'utilise pas pour filtrer les résultats.

Slack

Slack est une application de communication d'entreprise qui permet aux utilisateurs d'envoyer des messages et des pièces jointes via différents canaux publics et privés. Vous pouvez l'utiliser Amazon Kendra pour indexer vos chaînes publiques et privées Slack, envoyer et archiver des messages, des

fichiers et des pièces jointes, des messages directs et de groupe. Vous pouvez également choisir un contenu spécifique à filtrer.

> **Note**
>
> Amazon Kendra prend désormais en charge un connecteur Slack amélioré.
> La console a été automatiquement mise à niveau pour vous. Tous les nouveaux connecteurs que vous créez dans la console utiliseront l'architecture mise à niveau. Si vous utilisez l'API, vous devez désormais utiliser le TemplateConfigurationobjet au lieu de l'`SlackConfiguration`objet pour configurer votre connecteur.
> Les connecteurs configurés à l'aide de l'ancienne console et de l'ancienne architecture d'API continueront de fonctionner tels qu'ils ont été configurés. Toutefois, vous ne pourrez ni les modifier ni les mettre à jour. Si vous souhaitez modifier ou mettre à jour la configuration de votre connecteur, vous devez créer un nouveau connecteur.
> Nous vous recommandons de migrer le flux de travail de votre connecteur vers la version mise à niveau. Support pour les connecteurs configurés à l'aide de l'ancienne architecture devrait prendre fin en juin 2024.

Vous pouvez vous connecter Amazon Kendra à votre source de données Slack à l'aide de la Amazon Kendra console ou de l'TemplateConfigurationAPI.

Pour résoudre les problèmes liés à votre connecteur de source de données Amazon Kendra Slack, consultezDépannage des sources de données.

Rubriques

- Fonctionnalités prises en charge
- Prérequis
- Instructions de connexion
- En savoir plus

Fonctionnalités prises en charge

Amazon Kendra Le connecteur de source de données Slack prend en charge les fonctionnalités suivantes :

- Mappages de champs

- Contrôle d'accès utilisateur
- Filtres d'inclusion/exclusion
- Synchronisation complète et incrémentielle du contenu
- Cloud privé virtuel (VPC)

Prérequis

Avant de pouvoir Amazon Kendra indexer votre source de données Slack, apportez ces modifications à votre compte Slack et AWS à vos comptes.

Dans Slack, assurez-vous d'avoir :

- Vous avez configuré un OAuth jeton utilisateur Slack Bot ou un jeton utilisateur OAuth Slack. Vous pouvez choisir l'un ou l'autre des jetons pour vous connecter Amazon Kendra à votre source de données Slack. Un jeton est requis comme identifiant d'authentification. Consultez la documentation de Slack sur les jetons d'accès pour plus d'informations.

  > **Note**
  >
  > Si vous utilisez le jeton de bot dans vos informations d'identification Slack, vous ne pouvez pas indexer les messages directs et les messages de groupe et vous devez ajouter le jeton de bot à la chaîne que vous souhaitez indexer.

  > **Note**
  >
  > Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

- Vous avez noté votre identifiant d'équipe Slack sur l'URL de la page principale de votre espace de travail Slack. Par exemple, *https://app.slack.com/client/T0123456789/...* où se *T0123456789* trouve l'identifiant de l'équipe ?
- Les étendues et autorisations Oauth suivantes ont été ajoutées :

Portée du jeton utilisateur	Portée du jeton bot
• chaînes : historique • chaînes : lire • emoji : lire • fichiers:lire • groupes : historique • groupes : lire • im : histoire • im : lire • mpim : histoire • mpim : lire • team:read • users.profile : read • utilisateurs:read • utilisateurs:read.email	• chaînes : historique • canaux:gérer • chaînes : lire • conversations.connect : gérer • conversations.connect : lire • fichiers:lire • groupes : historique • groupes : lire • im : histoire • im : lire • mpim : histoire • mpim : lire • réactions : lire • team:read • groupes d'utilisateurs : lire • users.profile : read • utilisateurs:read • utilisateurs:read.email

- Il est vérifié que chaque document est unique dans Slack et dans les autres sources de données que vous prévoyez d'utiliser pour le même index. Chaque source de données que vous souhaitez utiliser pour un index ne doit pas contenir le même document dans toutes les sources de données. IDs Les documents sont globaux par rapport à un index et doivent être uniques par index.

Dans votre Compte AWS, assurez-vous d'avoir :

- Création d'un Amazon Kendra index et, si vous utilisez l'API, notez l'ID de l'index.
- Vous avez créé un IAM rôle pour votre source de données et, si vous utilisez l'API, notez l'ARN du IAM rôle.

> **Note**
>
> Si vous modifiez votre type d'authentification et vos informations d'identification, vous devez mettre à jour votre IAM rôle pour accéder au bon identifiant AWS Secrets Manager secret.

- Stockez vos informations d'authentification Slack dans un AWS Secrets Manager secret et, si vous utilisez l'API, notez l'ARN du secret.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

Si vous n'avez pas de IAM rôle ou de secret existant, vous pouvez utiliser la console pour créer un nouveau IAM rôle et un nouveau Secrets Manager secret lorsque vous connectez votre source de données Slack à Amazon Kendra. Si vous utilisez l'API, vous devez fournir l'ARN d'un IAM rôle et d'un Secrets Manager secret existants, ainsi qu'un identifiant d'index.

Instructions de connexion

Pour vous connecter Amazon Kendra à votre source de données Slack, vous devez fournir les informations nécessaires sur votre source de données Slack afin de Amazon Kendra pouvoir accéder à vos données. Si vous n'avez pas encore configuré Slack pour Amazon Kendra, consultezPrérequis.

Console

Pour vous connecter Amazon Kendra à Slack

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la.
2. Dans le volet de navigation de gauche, choisissez Index, puis choisissez l'index que vous souhaitez utiliser dans la liste des index.

> **Note**
>
> Vous pouvez choisir de configurer ou de modifier vos paramètres de contrôle d'accès utilisateur dans les paramètres de l'index.

3. Sur la page de démarrage, choisissez Ajouter une source de données.
4. Sur la page Ajouter une source de données, choisissez le connecteur Slack, puis sélectionnez Ajouter un connecteur. Si vous utilisez la version 2 (le cas échéant), choisissez le connecteur Slack avec le tag « V2.0 ».
5. Sur la page Spécifier les détails de la source de données, entrez les informations suivantes :
   a. Dans Nom et description, pour Nom de la source de données : entrez le nom de votre source de données. Vous pouvez inclure des tirets, mais pas des espaces.
   b. (Facultatif) Description : entrez une description facultative pour votre source de données.
   c. Dans la langue par défaut : choisissez une langue pour filtrer vos documents pour l'index. Sauf indication contraire, la langue par défaut est l'anglais. La langue spécifiée dans les métadonnées du document remplace la langue sélectionnée.
   d. Dans Balises, pour Ajouter une nouvelle balise : incluez des balises facultatives pour rechercher et filtrer vos ressources ou suivre vos AWS coûts.
   e. Choisissez Suivant.
6. Sur la page Définir l'accès et la sécurité, entrez les informations suivantes :
   a. Pour l'identifiant d'équipe de l'espace de travail Slack : l'identifiant d'équipe de votre espace de travail Slack. Vous trouverez votre identifiant d'équipe sur l'URL de la page principale de votre espace de travail Slack. Par exemple, *`https://app.slack.com/client/T0123456789/...`* où se *`T0123456789`* trouve l'identifiant de l'équipe ?
   b. Autorisation : activez ou désactivez les informations de la liste de contrôle d'accès (ACL) pour vos documents, si vous disposez d'une ACL et que vous souhaitez l'utiliser pour le contrôle d'accès. L'ACL indique les documents auxquels les utilisateurs et les groupes peuvent accéder. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.
   c. AWS Secrets Manager secret —Choisissez un secret existant ou créez-en un nouveau Secrets Manager pour stocker vos informations d'authentification Slack. Si vous

choisissez de créer un nouveau secret, une fenêtre AWS Secrets Manager secrète s'ouvre.

i. Entrez les informations suivantes dans la fenêtre Créer un AWS Secrets Manager secret :

A. Nom secret : nom de votre secret. Le préfixe « AmazonKendra -Slack » est automatiquement ajouté à votre nom secret.

B. Pour le jeton Slack : entrez les valeurs des informations d'authentification que vous avez configurées dans Slack.

ii. Enregistrez et ajoutez votre secret.

d. Virtual Private Cloud (VPC) —Vous pouvez choisir d'utiliser un VPC. Dans ce cas, vous devez ajouter des sous-réseaux et des groupes de sécurité VPC.

e. Identity Crawler : spécifiez s'il faut activer l'explorateur Amazon Kendra d'identité. Le robot d'exploration d'identité utilise les informations de la liste de contrôle d'accès (ACL) de vos documents pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Si vous disposez d'une ACL pour vos documents et que vous choisissez de l'utiliser, vous pouvez également choisir d'activer le robot d'exploration Amazon Kendra d'identité pour configurer le filtrage des résultats de recherche par contexte utilisateur. Dans le cas contraire, si Identity Crawler est désactivé, tous les documents peuvent faire l'objet d'une recherche publique. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'exploration d'identité est désactivé, vous pouvez également utiliser l'PutPrincipalMappingAPI pour télécharger les informations d'accès des utilisateurs et des groupes afin de filtrer le contexte utilisateur.

f. IAM rôle —Choisissez un IAM rôle existant ou créez-en un nouveau IAM pour accéder aux informations d'identification de votre référentiel et indexer le contenu.

> **Note**
>
> IAM les rôles utilisés pour les index ne peuvent pas être utilisés pour les sources de données. Si vous ne savez pas si un rôle existant est utilisé pour un index ou une FAQ, choisissez Créer un nouveau rôle pour éviter les erreurs.

g. Choisissez Suivant.

7. Sur la page Configurer les paramètres de synchronisation, entrez les informations suivantes :

a. Sélectionnez le type de contenu : sélectionnez les entités ou les types de contenu Slack que vous souhaitez analyser. Vous pouvez choisir parmi toutes les chaînes, les chaînes publiques, les chaînes privées, les messages de groupe et les messages privés.

b. Sélectionnez la date de début de l'analyse : entrez la date à laquelle vous souhaitez commencer à analyser votre contenu.

c. Pour une configuration supplémentaire : choisissez d'inclure les messages du bot et les messages archivés et d'utiliser des modèles d'expressions régulières pour inclure ou exclure certains contenus.

> **Note**
>
> Si vous choisissez d'inclure à la fois les noms des chaînes IDs et des chaînes, le connecteur Amazon Kendra Slack donnera la priorité aux chaînes IDs par rapport aux noms de chaînes.
> Si vous avez choisi d'inclure certains messages privés et de groupe, le connecteur Amazon Kendra Slack ignorera tous les messages privés et de groupe et n'analysera que les messages privés et de groupe que vous spécifiez.

d. Mode de synchronisation : choisissez la manière dont vous souhaitez mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation.

- Synchronisation complète : Indexez tout le contenu fraîchement, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
- Synchronisation nouvelle, modifiée ou supprimée : Indexez uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

e. Dans Calendrier d'exécution de la synchronisation, pour Fréquence : choisissez la fréquence à laquelle vous souhaitez synchroniser le contenu de votre source de données et mettre à jour votre index.

f. Choisissez Suivant.

8. Sur la page Définir les mappages de champs, entrez les informations suivantes :

   a. Champs de source de données par défaut : sélectionnez parmi les champs de source de données par défaut Amazon Kendra générés que vous souhaitez mapper à votre index.

   b. Ajouter un champ —Pour ajouter des champs de source de données personnalisés afin de créer un nom de champ d'index à mapper et le type de données du champ.

   c. Choisissez Suivant.

9. Sur la page Réviser et créer, vérifiez que les informations saisies sont correctes, puis sélectionnez Ajouter une source de données. Vous pouvez également choisir de modifier vos informations depuis cette page. Votre source de données apparaîtra sur la page Sources de données une fois que la source de données aura été ajoutée avec succès.

API

Pour vous connecter Amazon Kendra à Slack

Vous devez spécifier un code JSON du schéma de source de données à l'aide du TemplateConfigurationAPI. Vous devez fournir les informations suivantes :

- Source de données —Spécifiez le type de source de données comme `SLACK` lorsque vous utilisez TemplateConfigurationSchéma JSON. Spécifiez également la source de données comme `TEMPLATE` lorsque vous appelez le CreateDataSourceAPI.
- ID d'équipe Slack Workspace : identifiant d'équipe Slack que vous avez copié depuis l'URL de votre page principale de Slack.
- Depuis la date : date à laquelle vous avez commencé à analyser vos données auprès de votre équipe Slack Workspace. La date doit suivre ce format : yyyy-mm-dd.
- Mode de synchronisation : spécifiez comment Amazon Kendra mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation. Vous pouvez choisir entre :
  - `FORCED_FULL_CRAWL`pour indexer à nouveau tout le contenu, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.

  - `FULL_CRAWL`pour indexer uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
  - `CHANGE_LOG`pour indexer uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
- Identity Crawler : spécifiez s'il faut activer l'explorateur Amazon Kendra d'identité. Le robot d'exploration d'identité utilise les informations de la liste de contrôle d'accès (ACL) de vos documents pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Si vous disposez d'une ACL pour vos documents et que vous choisissez de l'utiliser, vous pouvez également choisir d'activer le robot d'exploration Amazon Kendra d'identité pour configurer le filtrage des résultats de recherche par contexte utilisateur. Dans le cas contraire, si Identity Crawler est désactivé, tous les documents peuvent faire l'objet d'une recherche publique. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'exploration d'identité est désactivé, vous pouvez également utiliser l'PutPrincipalMappingAPI pour télécharger les informations d'accès des utilisateurs et des groupes afin de filtrer le contexte utilisateur.
- Nom de ressource Amazon (ARN) secret : indiquez le nom de ressource Amazon (ARN) d'un Secrets Manager secret contenant les informations d'authentification de votre compte Slack. Le secret est stocké dans une structure JSON avec les clés suivantes :

```
{
    "slackToken": "token"
}
```

- IAM role —Spécifiez à quel `RoleArn` moment vous appelez `CreateDataSource` pour fournir à un IAM rôle l'autorisation d'accéder à votre Secrets Manager secret et d'appeler le public requis APIs pour le connecteur Slack et. Amazon Kendra Pour plus d'informations, consultez la section IAM Rôles des sources de données Slack.

Vous pouvez également ajouter les fonctionnalités optionnelles suivantes :

- Virtual Private Cloud (VPC) **`VpcConfiguration`** —Spécifiez le moment de votre appel. `CreateDataSource` Pour de plus amples informations, veuillez consulter Configuration Amazon Kendra pour utiliser un Amazon VPC.
- Chaînes spécifiques : filtrez par chaînes publiques ou privées et spécifiez certaines chaînes en fonction de leur identifiant.
- Types de chaînes et de messages : indiquez si vous Amazon Kendra devez indexer vos chaînes publiques et privées, vos messages de groupe et directs, ainsi que votre bot et vos messages archivés. Si vous utilisez un jeton de bot dans le cadre de vos identifiants d'authentification Slack, vous devez ajouter le jeton de bot à la chaîne que vous souhaitez indexer. Vous ne pouvez pas indexer les messages directs et les messages de groupe à l'aide d'un jeton de bot.
- Rétrospective : vous pouvez choisir de configurer un `lookBack` paramètre afin que le connecteur Slack analyse le contenu mis à jour ou supprimé jusqu'à un certain nombre d'heures avant la dernière synchronisation de votre connecteur.
- Filtres d'inclusion et d'exclusion : indiquez si vous souhaitez inclure ou exclure certains contenus de Slack. Si vous utilisez un jeton de bot dans le cadre de vos identifiants d'authentification Slack, vous devez ajouter le jeton de bot à la chaîne que vous souhaitez indexer. Vous ne pouvez pas indexer les messages directs et les messages de groupe à l'aide d'un jeton de bot.

  > **Note**
  >
  > La plupart des sources de données utilisent des modèles d'expressions régulières, qui sont des modèles d'inclusion ou d'exclusion appelés filtres. Si vous spécifiez un filtre d'inclusion, seul le contenu correspondant au filtre d'inclusion est indexé. Tout document qui ne correspond pas au filtre d'inclusion n'est pas indexé. Si vous spécifiez un filtre d'inclusion et d'exclusion, les documents correspondant au filtre d'exclusion ne sont pas indexés, même s'ils correspondent au filtre d'inclusion.

- Mappages de champs : choisissez de mapper les champs de votre source de données Slack à vos Amazon Kendra champs d'index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.

> **Note**
>
> Le champ du corps du document ou l'équivalent du corps du document pour vos documents est requis pour Amazon Kendra effectuer une recherche dans vos documents. Vous devez associer le nom du champ du corps du document dans votre source de données au nom du champ d'index`_document_body`. Tous les autres champs sont facultatifs.

Pour obtenir la liste des autres clés JSON importantes à configurer, voir Slack schéma de modèle.

En savoir plus

Pour en savoir plus sur l'intégration Amazon Kendra à votre source de données Slack, consultez :

- Exploitez les connaissances des espaces de travail Slack grâce à la recherche intelligente à l'aide du connecteur Slack Amazon Kendra

Zendesk

Zendesk est un système de gestion de la relation client qui aide les entreprises à automatiser et à améliorer les interactions avec le service client. Vous pouvez l'utiliser Amazon Kendra pour indexer vos tickets d'assistance Zendesk, les commentaires sur les tickets, les pièces jointes aux tickets, les articles du centre d'aide, les commentaires d'articles, les pièces jointes aux commentaires d'articles, les sujets de la communauté des guides, les publications communautaires et les commentaires des publications communautaires.

Vous pouvez filtrer par nom d'organisation si vous souhaitez indexer les tickets qui concernent uniquement une organisation spécifique. Vous pouvez également choisir de définir une date d'exploration à laquelle vous souhaitez commencer à explorer les données de Zendesk.

Vous pouvez vous connecter Amazon Kendra à votre source de données Zendesk à l'aide de la Amazon Kendra console et de l'TemplateConfigurationAPI.

Pour résoudre les problèmes liés à votre connecteur de source de données Amazon Kendra Zendesk, consultezDépannage des sources de données.

Rubriques

- Fonctionnalités prises en charge
- Prérequis
- Instructions de connexion
- En savoir plus
- Remarques

Fonctionnalités prises en charge

Amazon Kendra Le connecteur de source de données Zendesk prend en charge les fonctionnalités suivantes :

- Mappages de champs
- Contrôle d'accès utilisateur
- Filtres d'inclusion/exclusion
- Journal des modifications, synchronisation complète et incrémentielle du contenu
- Cloud privé virtuel (VPC)

Prérequis

Avant de pouvoir l'utiliser Amazon Kendra pour indexer votre source de données Zendesk, apportez ces modifications à votre Zendesk et AWS à vos comptes.

Dans Zendesk, assurez-vous d'avoir :

- Création d'un compte administratif Zendesk Suite (Professional/Enterprise).
- Vous avez noté l'URL de votre hôte Zendesk. Par exemple, *`https://{sub-domain}.zendesk.com/`*.

> **Note**
>
> (Sur place/sur serveur) Amazon Kendra vérifie si les informations de point de terminaison incluses sont les mêmes AWS Secrets Manager que celles spécifiées dans les détails de configuration de votre source de données. Cela permet de se protéger contre le problème de confusion des adjoints, qui est un problème de sécurité lorsqu'un utilisateur n'est pas autorisé à effectuer une action mais l'utilise Amazon Kendra comme proxy pour accéder au secret configuré et exécuter l'action. Si vous modifiez ultérieurement les informations

de votre point de terminaison, vous devez créer un nouveau secret pour synchroniser ces informations.

- Configurer l'authentification 0Auth 2.0 :
  1. Dans le Centre d'administration, accédez à Applications et intégrations > APIs > API Zendesk.
  2. Sélectionnez l'onglet OAuth Clients et cliquez sur « Ajouter un OAuth client ».
  3. Configurez les informations suivantes sur le OAuth client : définissez le nom et la description du client, définissez le type de client sur « Confidentiel », ajoutez la redirection appropriée URLs, enregistrez et stockez en toute sécurité le secret client généré.
  4. Assurez-vous que le OAuth client dispose de l'étendue « lecture » requise (ou « lecture-écriture » si vous avez besoin d'un accès en écriture).
  5. Générez un jeton d'accès à l'aide du flux de subvention implicite :
     - Dans un navigateur, accédez à : `https://{subdomain}.zendesk.com/oauth/authorizations/new?response_type=token&client_id=<your_client_id>&scope=read`
     - Authentifiez et autorisez l'application lorsque vous y êtes invité.
     - Après autorisation, vous serez redirigé vers une URL contenant le jeton d'accès dans le fragment (après le #). Extrayez ce jeton.
  6. Stockez le jeton d'accès généré en toute sécurité. Il s'agit du « jeton de subvention implicite » que vous utiliserez pour l'intégration de Kendra.

- **Note**

  Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

- Facultatif : un certificat SSL a été installé Amazon Kendra pour permettre la connexion.
- Il est vérifié que chaque document est unique dans Zendesk et dans les autres sources de données que vous prévoyez d'utiliser pour le même index. Chaque source de données que vous souhaitez utiliser pour un index ne doit pas contenir le même document dans toutes les sources de données. IDs Les documents sont globaux par rapport à un index et doivent être uniques par index.

Dans votre Compte AWS, assurez-vous d'avoir :

- Création d'un Amazon Kendra index et, si vous utilisez l'API, notez l'ID de l'index.
- Vous avez créé un IAM rôle pour votre source de données et, si vous utilisez l'API, notez l'ARN du IAM rôle.

> **Note**
>
> Si vous modifiez votre type d'authentification et vos informations d'identification, vous devez mettre à jour votre IAM rôle pour accéder au bon identifiant AWS Secrets Manager secret.

- Stockez vos informations d'authentification Zendesk dans un AWS Secrets Manager secret et, si vous utilisez l'API, notez l'ARN du secret.

> **Note**
>
> Nous vous recommandons d'actualiser ou de modifier régulièrement vos informations d'identification et votre code secret. Fournissez uniquement le niveau d'accès nécessaire pour votre propre sécurité. Nous vous déconseillons de réutiliser les informations d'identification et les secrets entre les sources de données et les versions 1.0 et 2.0 du connecteur (le cas échéant).

Si vous n'avez pas de IAM rôle ou de secret existant, vous pouvez utiliser la console pour créer un nouveau IAM rôle et un nouveau Secrets Manager secret lorsque vous connectez votre source de données Zendesk à Amazon Kendra. Si vous utilisez l'API, vous devez fournir l'ARN d'un IAM rôle et d'un Secrets Manager secret existants, ainsi qu'un identifiant d'index.

Instructions de connexion

Pour vous connecter Amazon Kendra à votre source de données Zendesk, vous devez fournir les informations nécessaires sur votre source de données Zendesk afin de Amazon Kendra pouvoir accéder à vos données. Si vous n'avez pas encore configuré Zendesk pour Amazon Kendra, consultezPrérequis.

Console

Pour vous connecter Amazon Kendra à Zendesk

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la.
2. Dans le volet de navigation de gauche, choisissez Index, puis choisissez l'index que vous souhaitez utiliser dans la liste des index.

> **Note**
>
> Vous pouvez choisir de configurer ou de modifier vos paramètres de contrôle d'accès utilisateur dans les paramètres de l'index.

3. Sur la page de démarrage, choisissez Ajouter une source de données.
4. Sur la page Ajouter une source de données, choisissez le connecteur Zendesk, puis sélectionnez Ajouter un connecteur. Si vous utilisez la version 2 (le cas échéant), choisissez le connecteur Zendesk avec la balise « V2.0 ».
5. Sur la page Spécifier les détails de la source de données, entrez les informations suivantes :
   a. Dans Nom et description, pour Nom de la source de données : entrez le nom de votre source de données. Vous pouvez inclure des tirets, mais pas des espaces.
   b. (Facultatif) Description : entrez une description facultative pour votre source de données.
   c. Dans la langue par défaut : choisissez une langue pour filtrer vos documents pour l'index. Sauf indication contraire, la langue par défaut est l'anglais. La langue spécifiée dans les métadonnées du document remplace la langue sélectionnée.
   d. Dans Balises, pour Ajouter une nouvelle balise : incluez des balises facultatives pour rechercher et filtrer vos ressources ou suivre vos AWS coûts.
   e. Choisissez Suivant.
6. Sur la page Définir l'accès et la sécurité, entrez les informations suivantes :
   a. URL Zendesk —Entrez votre URL Zendesk. Par exemple, *https://{sub-domain}.zendesk.com/*.
   b. Autorisation : activez ou désactivez les informations de la liste de contrôle d'accès (ACL) pour vos documents, si vous disposez d'une ACL et que vous souhaitez l'utiliser pour le contrôle d'accès. L'ACL indique les documents auxquels les utilisateurs et les groupes peuvent accéder. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.
   c. AWS Secrets Manager secret —Choisissez un secret existant ou créez-en un nouveau Secrets Manager pour stocker vos informations d'authentification Zendesk. Si vous

choisissez de créer un nouveau secret, une fenêtre AWS Secrets Manager secrète s'ouvre.

i. Créez un nouveau secret avec la structure suivante :

```
{
        "hostUrl": "https://yoursubdomain.zendesk.com/",
        "accessToken": "your_implicit_grant_access_token"
}
```

ii. Enregistrez et ajoutez votre secret.

d. Virtual Private Cloud (VPC) —Vous pouvez choisir d'utiliser un VPC. Dans ce cas, vous devez ajouter des sous-réseaux et des groupes de sécurité VPC.

e. Identity Crawler : spécifiez s'il faut activer l'explorateur Amazon Kendra d'identité. Le robot d'exploration d'identité utilise les informations de la liste de contrôle d'accès (ACL) de vos documents pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Si vous disposez d'une ACL pour vos documents et que vous choisissez de l'utiliser, vous pouvez également choisir d'activer le robot d'exploration Amazon Kendra d'identité pour configurer le filtrage des résultats de recherche par contexte utilisateur. Sinon, si le robot d'identification est désactivé, tous les documents peuvent être consultés publiquement. Si vous souhaitez utiliser le contrôle d'accès pour vos documents et que le robot d'exploration d'identité est désactivé, vous pouvez également utiliser l'PutPrincipalMappingAPI pour télécharger les informations d'accès des utilisateurs et des groupes afin de filtrer le contexte utilisateur.

f. IAM rôle —Choisissez un IAM rôle existant ou créez-en un nouveau IAM pour accéder aux informations d'identification de votre référentiel et indexer le contenu.

> **Note**
>
> IAM les rôles utilisés pour les index ne peuvent pas être utilisés pour les sources de données. Si vous ne savez pas si un rôle existant est utilisé pour un index ou une FAQ, choisissez Créer un nouveau rôle pour éviter les erreurs.

g. Choisissez Suivant.

7. Sur la page Configurer les paramètres de synchronisation, entrez les informations suivantes :

a. Sélectionnez le contenu : sélectionnez les types de contenu que vous souhaitez explorer à partir des tickets, des articles du centre d'aide, des sujets communautaires, etc.

b. Nom de l'organisation : entrez les noms des organisations Zendesk pour filtrer le contenu.

c. Date de début de synchronisation : entrez la date à partir de laquelle vous souhaitez commencer à analyser votre contenu.

d. Modèles Regex : ajoutez des modèles d'expressions régulières pour inclure ou exclure certains fichiers. Vous pouvez ajouter jusqu'à 100 motifs.

e. Mode de synchronisation : choisissez la manière dont vous souhaitez mettre à jour votre index lorsque le contenu de votre source de données change. Lorsque vous synchronisez votre source de données Amazon Kendra pour la première fois, tout le contenu est analysé et indexé par défaut. Vous devez exécuter une synchronisation complète de vos données en cas d'échec de la synchronisation initiale, même si vous ne choisissez pas l'option de synchronisation complète comme mode de synchronisation.

   - Synchronisation complète : Indexez tout le contenu fraîchement, en remplaçant le contenu existant chaque fois que votre source de données se synchronise avec votre index.
   - Nouvelle synchronisation modifiée : Indexez uniquement le contenu nouveau et modifié chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.
   - Synchronisation nouvelle, modifiée ou supprimée : Indexez uniquement le contenu nouveau, modifié et supprimé chaque fois que votre source de données se synchronise avec votre index. Amazon Kendra peut utiliser le mécanisme de votre source de données pour suivre les modifications de contenu et indexer le contenu modifié depuis la dernière synchronisation.

f. Dans Synchroniser le calendrier d'exécution pour la fréquence : choisissez la fréquence à laquelle vous souhaitez synchroniser le contenu de votre source de données et mettre à jour votre index.

g. Choisissez Suivant.

8. Sur la page Définir les mappages de champs, entrez les informations suivantes :

   a. Champs de source de données par défaut : sélectionnez parmi les champs de source de données par défaut Amazon Kendra générés que vous souhaitez mapper à votre index.

b. Ajouter un champ —Pour ajouter des champs de source de données personnalisés afin de créer un nom de champ d'index à mapper et le type de données du champ.

c. Choisissez Suivant.

9. Sur la page Réviser et créer, vérifiez que les informations saisies sont correctes, puis sélectionnez Ajouter une source de données. Vous pouvez également choisir de modifier vos informations depuis cette page. Votre source de données apparaîtra sur la page Sources de données une fois que la source de données aura été ajoutée avec succès.

API

Pour vous connecter Amazon Kendra à Zendesk

Vous devez spécifier un JSON du schéma de source de données à l'aide de l'TemplateConfigurationAPI. Vous devez fournir les informations suivantes :

- Source de données —Spécifiez le type de source de données tel que `ZENDESK` lorsque vous utilisez le schéma TemplateConfigurationJSON. Spécifiez également la source de données `TEMPLATE` lorsque vous appelez l'CreateDataSourceAPI.
- URL de l'hôte : fournissez l'URL de votre hôte Zendesk dans le cadre de la configuration de la connexion ou des détails du point de terminaison du référentiel. Par exemple, *`https://yoursubdomain.zendesk.com`*.
- Journal des modifications : Amazon Kendra faut-il utiliser le mécanisme du journal des modifications des sources de données Zendesk pour déterminer si un document doit être mis à jour dans l'index.

  > **Note**
  >
  > Utilisez le journal des modifications si vous ne Amazon Kendra souhaitez pas numériser tous les documents. Si votre journal des modifications est volumineux, la numérisation des documents de la source de données Zendesk peut prendre Amazon Kendra moins de temps que le traitement du journal des modifications. Si vous synchronisez votre source de données Zendesk avec votre index pour la première fois, tous les documents sont numérisés.

- Nom de ressource Amazon secret (ARN) : fournissez le nom de ressource Amazon (ARN) d'un Secrets Manager secret contenant les informations d'authentification de votre compte Zendesk. Le secret est stocké dans une structure JSON avec les clés suivantes :

```
{
    "hostUrl": "https://yoursubdomain.zendesk.com",
    "clientId": "client ID",
    "clientSecret": "Zendesk client secret",
    "userName": "Zendesk user name",
    "password": "Zendesk password"
}
```

- IAM role —Spécifiez à quel `RoleArn` moment vous appelez `CreateDataSource` pour fournir à un IAM rôle les autorisations d'accéder à votre Secrets Manager secret et d'appeler le public requis APIs pour le connecteur Zendesk et. Amazon Kendra Pour plus d'informations, consultez la section IAM Rôles pour les sources de données Zendesk.

Vous pouvez également ajouter les fonctionnalités optionnelles suivantes :

- Virtual Private Cloud (VPC) **`VpcConfiguration`** —Spécifiez le moment de votre appel. `CreateDataSource` Pour de plus amples informations, veuillez consulter Configuration Amazon Kendra pour utiliser un Amazon VPC.
- Types de documents/contenus : spécifiez si vous souhaitez explorer :
  - Support : tickets, commentaires sur les tickets, pièces jointes aux commentaires sur les and/or tickets
  - Articles du centre d'aide, pièces jointes aux articles et commentaires d'articles
  - Guidez les sujets, les publications ou les commentaires de la communauté
- Filtres d'inclusion et d'exclusion : indiquez si vous souhaitez inclure ou exclure certains contenus de Slack. Si vous utilisez un jeton de bot dans le cadre de vos identifiants d'authentification Slack, vous devez ajouter le jeton de bot à la chaîne que vous souhaitez indexer. Vous ne pouvez pas indexer les messages directs et les messages de groupe à l'aide d'un jeton de bot.

> **Note**
>
> La plupart des sources de données utilisent des modèles d'expressions régulières, qui sont des modèles d'inclusion ou d'exclusion appelés filtres. Si vous spécifiez un filtre d'inclusion, seul le contenu correspondant au filtre d'inclusion est indexé. Tout document qui ne correspond pas au filtre d'inclusion n'est pas indexé. Si vous spécifiez

un filtre d'inclusion et d'exclusion, les documents correspondant au filtre d'exclusion ne sont pas indexés, même s'ils correspondent au filtre d'inclusion.

- Filtrage du contexte utilisateur et contrôle d'accèsAmazon Kendra : analyse la liste de contrôle d'accès (ACL) de vos documents, si vous disposez d'une ACL pour vos documents. Les informations ACL sont utilisées pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Pour plus d'informations, consultez la section Filtrage du contexte utilisateur.
- Mappages de champs : choisissez de mapper les champs de votre source de données Zendesk à vos Amazon Kendra champs d'index. Pour plus d'informations, veuillez consulter la rubrique Mappage des champs de source de données.

> **Note**
>
> Le champ du corps du document ou l'équivalent du corps du document pour vos documents est requis pour Amazon Kendra effectuer une recherche dans vos documents. Vous devez associer le nom du champ du corps du document dans votre source de données au nom du champ d'index`_document_body`. Tous les autres champs sont facultatifs.

Pour obtenir la liste des autres clés JSON importantes à configurer, consultez le schéma du modèle Zendesk.

En savoir plus

Pour en savoir plus sur l'intégration Amazon Kendra à votre source de données Zendesk, consultez :

- Découvrez les informations de Zendesk grâce à la recherche Amazon Kendra intelligente

Remarques

- Lorsque les listes de contrôle d'accès (ACLs) sont activées, l'option « Synchroniser uniquement le contenu nouveau ou modifié » n'est pas disponible en raison des limites de l'API Zendesk. Nous vous recommandons d'utiliser plutôt les modes « Synchronisation complète » ou « Synchronisation du contenu nouveau, modifié ou supprimé », ou de les désactiver ACLs si vous devez utiliser ce mode de synchronisation.

Cartographie des champs de source de données

Amazon Kendra les connecteurs de source de données peuvent mapper les champs de document ou de contenu de votre source de données aux champs de votre Amazon Kendra index. Par défaut, chaque connecteur est conçu pour analyser des champs de source de données spécifiques. Les champs de source de données par défaut et leurs propriétés ne peuvent pas être modifiés ou personnalisés. Sur la Amazon Kendra console, les champs par défaut et les propriétés des champs par défaut qui ne peuvent pas être modifiés sont grisés.

Amazon Kendra les connecteurs vous permettent également de mapper des champs de document ou de contenu personnalisés de votre source de données aux champs personnalisés de votre index. Par exemple, si votre source de données contient un champ appelé « département » qui contient les informations relatives au service d'un document, vous pouvez le mapper à un champ d'index appelé « Département ». Ainsi, vous pouvez utiliser le champ lorsque vous interrogez des documents.

Vous pouvez également mapper des champs Amazon Kendra réservés ou communs tels que`_created_at`. Si votre source de données possède un champ appelé « creation_date », vous pouvez le mapper au champ Amazon Kendra réservé équivalent appelé. `_created_at` Pour plus d'informations sur les champs Amazon Kendra réservés, consultez la section Attributs ou champs du document.

Vous pouvez mapper les champs de la plupart des sources de données. Vous pouvez créer des mappages de champs pour les sources de données suivantes :

- Gestionnaire d'expérience Adobe
- En plein air
- Aurora (MySQL)
- Aurora (PostgreSQL)
- Amazon FSx (Fenêtres)
- Amazon FSx (NetApp ONTAP)
- Amazon RDS/Aurora
- Amazon RDS (Microsoft SQL Server)
- Amazon RDS (MySQL)
- Amazon RDS (Oracle)
- Amazon RDS (PostgreSQL)
- Amazon Kendra robot d'exploration Web

- WorkDocs
- Box (Cube)
- Confluence
- Dropbox
- Drupal
- GitHub
- Disques durs Google Workspace
- Gmail
- IBM DB2
- Jira
- Microsoft Exchange
- Microsoft OneDrive
- Microsoft SharePoint
- Microsoft Teams
- Microsoft SQL Server
- Microsoft Yammer
- MySQL
- Oracle Database
- PostgreSQL
- Quip
- Salesforce
- ServiceNow
- Slack
- Zendesk

Si vous stockez vos documents dans un compartiment S3 ou une source de données S3, vous spécifiez vos champs à l'aide d'un fichier de métadonnées JSON. Pour plus d'informations, consultez la section Connecteur de source de données S3.

Le mappage des champs de votre source de données vers un champ d'index s'effectue en trois étapes :

1. Créez un index. Pour plus d'informations, consultez la section Création d'un index.
2. Mettez à jour l'index pour ajouter des champs.
3. Créez une source de données et incluez des mappages de champs pour mapper les champs réservés et tous les champs personnalisés pour Amazon Kendra indexer les champs.

Pour mettre à jour l'index afin d'ajouter des champs personnalisés, utilisez la console pour modifier les mappages de champs de la source de données et ajouter un champ personnalisé ou utilisez l'UpdateIndexAPI. Vous pouvez ajouter un total de 500 champs personnalisés à votre index.

Pour les sources de données de base de données, si le nom de la colonne de base de données correspond au nom d'un champ réservé, le champ et la colonne sont automatiquement mappés.

Avec l'UpdateIndexAPI, vous pouvez ajouter des champs réservés et personnalisés à l'aide de`DocumentMetadataConfigurationUpdates`.

L'exemple JSON suivant permet `DocumentMetadataConfigurationUpdates` d'ajouter un champ appelé « Department » à l'index.

```
"DocumentmetadataConfigurationUpdates": [
   {
       "Name": "Department",
       "Type": "STRING_VALUE"
   }
]
```

Lorsque vous créez le champ, vous avez la possibilité de définir la manière dont le champ est utilisé pour la recherche. Sélectionnez parmi les éléments suivants :

- Affichable —Détermine si le champ est renvoyé dans la réponse à la requête. L'argument par défaut est `true`.
- Facetable —Indique que le champ peut être utilisé pour créer des facettes. L'argument par défaut est `false`.
- Consultable —Détermine si le champ est utilisé dans la recherche. La valeur par défaut est `true` pour les champs de chaîne et `false` pour les champs de nombre et de date.
- Triable —Indique que le champ peut être utilisé pour trier la réponse d'une requête. Ne peut être défini que pour les champs de date, de numéro et de chaîne. Impossible de définir les champs de liste de chaînes.

L'exemple JSON suivant permet `DocumentMetadataConfigurationUpdates` d'ajouter un champ appelé « Department » à l'index et de le marquer comme facetable.

```
"DocumentMetadataConfigurationUpdates": [
   {
       "Name": "Department",
       "Type": "STRING_VALUE",
       "Search": {
           "Facetable": true
       }
   }
]
```

Utilisation de champs de document Amazon Kendra réservés ou communs

Avec l'UpdateIndex API, vous pouvez créer des champs réservés ou communs en utilisant `DocumentMetadataConfigurationUpdates` et en spécifiant le nom du champ d'index Amazon Kendra réservé à mapper à votre attribute/field nom de document équivalent. Vous pouvez également créer des champs personnalisés. Si vous utilisez un connecteur de source de données, la plupart incluent des mappages de champs qui font correspondre les champs de votre document de source de données aux champs d' Amazon Kendra index. Si vous utilisez la console, vous mettez à jour les champs en sélectionnant votre source de données, en sélectionnant l'action de modification, puis en passant à côté de la section des mappages de champs pour configurer la source de données.

Vous pouvez configurer l'`Search`objet pour définir un champ comme affichable, facetable, consultable ou triable. Vous pouvez configurer l'`Relevance`objet pour définir l'ordre de classement d'un champ, la durée d'augmentation ou la période à appliquer aux valeurs de renforcement, de fraîcheur, de valeur d'importance et d'importance mappées à des valeurs de champ spécifiques. Si vous utilisez la console, vous pouvez définir les paramètres de recherche d'un champ en sélectionnant l'option à facettes dans le menu de navigation. Pour définir le réglage de la pertinence, sélectionnez l'option permettant de rechercher votre index dans le menu de navigation, entrez une requête et utilisez les options du panneau latéral pour ajuster la pertinence de la recherche. Vous ne pouvez pas modifier le type de champ une fois que vous l'avez créé.

Amazon Kendra possède les champs de document réservés ou communs suivants que vous pouvez utiliser :

- `_authors`—Une liste d'un ou de plusieurs auteurs responsables du contenu du document.

- `_category`: catégorie qui place un document dans un groupe spécifique.
- `_created_at`: date et heure au format ISO 8601 auxquelles le document a été créé. Par exemple, 2012-03-25T12:30:10+01:00 est le format de date et d'heure ISO 8601 pour le 25 mars 2012 à 12 h 30 (plus 10 secondes) à l'heure d'Europe centrale.
- `_data_source_id`: identifiant de la source de données qui contient le document.
- `_document_body`: le contenu du document.
- `_document_id`—Un identifiant unique pour le document.
- `_document_title`: le titre du document.
- `_excerpt_page_number`: le numéro de page d'un fichier PDF où apparaît l'extrait du document. Si votre index a été créé avant le 8 septembre 2020, vous devez réindexer vos documents avant de pouvoir utiliser cet attribut.
- `_faq_id`—S'il s'agit d'un document de type question-réponse (FAQ), un identifiant unique pour la FAQ.
- `_file_type`: le type de fichier du document, tel que pdf ou doc.
- `_last_updated_at`: date et heure au format ISO 8601 auxquelles le document a été mis à jour pour la dernière fois. Par exemple, 2012-03-25T12:30:10+01:00 est le format de date et d'heure ISO 8601 pour le 25 mars 2012 à 12 h 30 (plus 10 secondes) à l'heure d'Europe centrale.
- `_source_uri`: l'URI où le document est disponible. Par exemple, l'URI du document sur le site Web d'une entreprise.
- `_version`—Identifiant pour la version spécifique d'un document.
- `_view_count`: le nombre de fois que le document a été consulté.
- `_language_code`(String) : code d'une langue qui s'applique au document. La valeur par défaut est l'anglais si vous ne spécifiez aucune langue. Pour plus d'informations sur les langues prises en charge, y compris leurs codes, voir Ajout de documents dans des langues autres que l'anglais.

Pour les champs personnalisés, vous pouvez les créer à l'`DocumentMetadataConfigurationUpdates`aide de l'`UpdateIndex`API, comme vous le faites lorsque vous créez un champ réservé ou commun. Vous devez définir le type de données approprié pour votre champ personnalisé. Si vous utilisez la console, vous mettez à jour les champs en sélectionnant votre source de données, en sélectionnant l'action de modification, puis en passant à côté de la section des mappages de champs pour configurer la source de données. Certaines sources de données ne prennent pas en charge l'ajout de nouveaux champs ou de champs personnalisés. Vous ne pouvez pas modifier le type de champ une fois que vous l'avez créé.

Les types que vous pouvez définir pour les champs personnalisés sont les suivants :

- Date
- Nombre
- Chaîne
- Liste de chaînes

Si vous avez ajouté des documents à l'index à l'aide fields/attributes de l'BatchPutDocumentAPI, `Attributes` répertorie vos documents et créez des champs à l'aide de l'`DocumentAttribute`objet.

Pour les documents indexés à partir d'une source de Amazon S3 données, vous créez des champs à l'aide d'un fichier de métadonnées JSON qui inclut les informations des champs.

Si vous utilisez une base de données prise en charge comme source de données, vous pouvez configurer vos champs à l'aide de l'option de mappage de champs.

Ajouter des documents dans des langues autres que l'anglais

Vous pouvez indexer des documents dans plusieurs langues. Si vous ne spécifiez aucune langue, Amazon Kendra indexe les documents en anglais par défaut. Vous incluez le code de langue d'un document dans les métadonnées du document sous forme de champ. Voir Mappages de champs et Attributs personnalisés pour plus d'informations sur le `_language_code` champ d'un document.

Vous pouvez spécifier le code de langue de tous vos documents dans votre source de données lorsque vous appelez CreateDataSource. Si aucun code de langue n'est spécifié dans un champ de métadonnées, le document est indexé à l'aide du code de langue spécifié pour tous les documents au niveau de la source de données. Dans la console, vous pouvez indexer des documents dans une langue prise en charge uniquement au niveau de la source de données. Accédez à Sources de données, puis à la page Spécifier les détails de la source de données, puis choisissez une langue dans la liste déroulante Langue.

Vous pouvez également rechercher ou interroger des documents dans une langue prise en charge. Pour plus d'informations, consultez la section Recherche dans les langues.

Les langues suivantes et leurs codes sont pris en charge (anglais ou `en` pris en charge par défaut si vous ne spécifiez aucune langue). Ce tableau inclut les langues qui prennent Amazon Kendra en charge la recherche sémantique complète, ainsi que les langues qui ne prennent en charge que la

correspondance de mots clés simples. Les langues qui prennent en charge la recherche sémantique complète sont signalées par un astérisque et apparaissent en gras dans le tableau suivant. L'anglais (langue par défaut) est également pris en charge avec la recherche sémantique complète.

Nom de langue	Code de langue
Arabe	`ar`
Arménien	`hy`
Basque	`eu`
Bengali	`bn`
Bulgare	`bg`
Catalan	`ca`
Chinois : simplifié et traditionnel*	`zh`
Tchèque	`cs`
Danois	`da`
Néerlandais	`nl`
Finnois	`fi`
Français — y compris le français (Canada) *	`fr`
Galicien	`gl`
Allemand*	`de`
Grec	`el`
Hindi	`hi`
Hongrois	`hu`
Indonésien	`id`

Nom de langue	Code de langue
Irlandais	`ga`
Italien	`it`
Japonais*	`ja`
Coréen*	`ko`
Letton	`lv`
Lituanien	`lt`
Norvégien	`no`
Persan	`fa`
Portugais	`pt`
Portugais (Brésil) *	`pt-BR`
Roumain	`ro`
Russe	`ru`
Sorani	`ckb`
Espagnol — inclut l'espagnol (Mexique) *	`es`
Suédois	`sv`
Turc	`tr`

*La recherche sémantique est prise en charge pour la langue.

Pour les langues qui prennent en charge la recherche sémantique, les fonctionnalités suivantes sont prises en charge.

- La pertinence du document va au-delà de la simple correspondance de mots clés.
- FAQs au-delà de la simple correspondance de mots clés.

- Extraire des réponses à partir de documents en fonction Amazon Kendra de sa compréhension écrite.
- Tranches de confiance (très élevée, élevée, moyenne et faible) des résultats de recherche.

Pour les langues qui ne prennent pas en charge la recherche sémantique, la simple correspondance de mots clés est prise en charge pour la pertinence du document et FAQs.

Les synonymes (y compris les synonymes personnalisés), l'apprentissage progressif et les commentaires, ainsi que les suggestions de requêtes ne sont pris en charge qu'en anglais (langue par défaut).

Configuration Amazon Kendra pour utiliser un Amazon VPC

Amazon Kendra peut se connecter à un cloud privé virtuel (VPC) que vous avez créé Amazon Virtual Private Cloud pour indexer le contenu stocké dans les sources de données exécutées dans votre cloud privé. Lorsque vous créez un connecteur de source de données, vous pouvez fournir des identifiants de groupe de sécurité et de sous-réseau pour le sous-réseau qui contient votre source de données. À partir de ces informations, il Amazon Kendra crée une interface réseau élastique qu'il utilise pour communiquer en toute sécurité avec votre source de données au sein de votre VPC.

Pour configurer un connecteur de source de Amazon Kendra données avec Amazon VPC, vous pouvez utiliser l'opération AWS Management Console ou l'opération CreateDataSourceAPI. Si vous utilisez la console, vous connectez un VPC pendant le processus de configuration du connecteur.

> **Note**
>
> Amazon VPC Cette fonctionnalité est facultative lors de la configuration d'un connecteur de source de Amazon Kendra données. Si votre source de données est accessible depuis l'Internet public, il n'est pas nécessaire d'activer Amazon VPC cette fonctionnalité. Les connecteurs de source Amazon Kendra de données ne sont pas tous compatibles Amazon VPC.

Si votre source de données n'est pas active Amazon VPC et n'est pas accessible depuis l'Internet public, vous devez d'abord la connecter à votre VPC via un réseau privé virtuel (VPN). Vous pouvez ensuite connecter votre source de données à Amazon Kendra l'aide d'une combinaison de Amazon VPC et AWS Virtual Private Network. Pour plus d'informations sur la configuration d'un VPN, consultez la AWS VPN documentation.

Rubriques

- Configuration de la Amazon VPC prise en charge des Amazon Kendra connecteurs
- Configuration d'une source Amazon Kendra de données à laquelle se connecter Amazon VPC
- Connexion à une base de données dans un VPC
- Résolution des problèmes de connexion VPC

Configuration de la Amazon VPC prise en charge des Amazon Kendra connecteurs

Amazon VPC Pour configurer en vue de l'utiliser avec vos Amazon Kendra connecteurs, procédez comme suit.

Étapes

- Étape 1. Créez des Amazon VPC sous-réseaux pour Amazon Kendra
- Étape 2. Créez des groupes Amazon VPC de sécurité pour Amazon Kendra
- Étape 3. Configurez votre source de données externe et Amazon VPC

Étape 1. Créez des Amazon VPC sous-réseaux pour Amazon Kendra

Créez ou choisissez un Amazon VPC sous-réseau existant qui Amazon Kendra peut être utilisé pour accéder à votre source de données. Les sous-réseaux préparés doivent se trouver dans l'une des zones Régions AWS de disponibilité suivantes :

- USA Ouest (Oregon) /us-west-2—usw2-az1, usw2-az2, usw2-az3
- USA Est (Virginie du Nord) /us-east-1—use1-az1, use1-az2, use1-az4
- USA Est (Ohio) /us-east-2—use2-az1, use2-az2, use2-az3
- Asie-Pacifique (Tokyo) /ap-northeast-1—apne1-az1, apne1-az2, apne1-az4
- Asie-Pacifique (Mumbai) /ap-south-1—aps1-az1, aps1-az2, aps1-az3
- Asie-Pacifique (Singapour) /ap-southeast-1—apse1-az1, apse1-az2, apse1-az3
- Asie-Pacifique (Sydney) /ap-southeast-2—apse2-az1, apse2-az2, apse2-az3
- Canada (Centre) /ca-central-1—cac1-az1, cac1-az2, cac1-az4
- Europe (Irlande) /eu-west-1—euw1-az1, uew1-az2, euw1-az3
- Europe (Londres) /eu-west-2—euw2-az1, euw2-az2, euw2-az3

Votre source de données doit être accessible depuis les sous-réseaux que vous avez fournis au Amazon Kendra connecteur.

Pour plus d'informations sur la configuration des Amazon VPC sous-réseaux, consultez la section Sous-réseaux correspondants Amazon VPC dans le guide de l'utilisateur Amazon VPC.

Si vous Amazon Kendra devez acheminer la connexion entre deux ou plusieurs sous-réseaux, vous pouvez préparer plusieurs sous-réseaux. Par exemple, le sous-réseau qui contient votre source de données n'a plus d'adresses IP. Dans ce cas, vous pouvez fournir Amazon Kendra un sous-réseau supplémentaire doté d'adresses IP suffisantes et connecté au premier sous-réseau. Si vous répertoriez plusieurs sous-réseaux, ceux-ci doivent pouvoir communiquer entre eux.

Étape 2. Créez des groupes Amazon VPC de sécurité pour Amazon Kendra

Pour connecter votre connecteur de source de Amazon Kendra données à Amazon VPC, vous devez préparer un ou plusieurs groupes de sécurité à partir de votre VPC auxquels vous souhaitez les attribuer. Amazon Kendra Les groupes de sécurité seront associés à l'interface elastic network créée par Amazon Kendra. Cette interface réseau contrôle le trafic entrant et sortant à destination et en provenance des sous-réseaux Amazon Kendra lors de l' Amazon VPC accès aux sous-réseaux.

Assurez-vous que les règles sortantes de votre groupe de sécurité autorisent le trafic provenant Amazon Kendra des connecteurs de source de données à accéder aux sous-réseaux et à la source de données avec lesquels vous allez effectuer la synchronisation. Par exemple, vous pouvez utiliser un MySQL connecteur pour effectuer une synchronisation à partir d'une MySQL base de données. Si vous utilisez le port par défaut, les groupes de sécurité doivent autoriser l'accès Amazon Kendra au port 3306 sur l'hôte qui exécute la base de données.

Nous vous recommandons de configurer un groupe de sécurité par défaut avec les valeurs suivantes Amazon Kendra à utiliser :

- Règles de trafic entrant — Si vous choisissez de laisser ce champ vide, tout le trafic entrant sera bloqué.
- Règles de trafic sortant : ajoutez une règle pour autoriser tout le trafic sortant afin de lancer les demandes de synchronisation à partir de votre source de données. Amazon Kendra
  - Version IP — IPv4
  - Type — Tout le trafic
  - Protocole — Tout le trafic
  - Gamme Port — Tous

- Destination : 0.0.0.0/0

Pour plus d'informations sur la configuration des groupes Amazon VPC de sécurité, consultez la section Règles des groupes de sécurité dans le guide de l'utilisateur Amazon VPC.

Étape 3. Configurez votre source de données externe et Amazon VPC

Assurez-vous que la configuration des autorisations et les paramètres réseau de votre source de données externe sont corrects Amazon Kendra pour y accéder. Vous trouverez des instructions détaillées sur la configuration de vos sources de données dans la section des prérequis de chaque page de connecteur.

Vérifiez également vos Amazon VPC paramètres et assurez-vous que votre source de données externe est accessible depuis le sous-réseau auquel vous allez l'attribuer Amazon Kendra. Pour ce faire, nous vous recommandons de créer une Amazon EC2 instance dans le même sous-réseau avec les mêmes groupes de sécurité et de tester l'accès à votre source de données à partir de cette Amazon EC2 instance. Pour plus d'informations, consultez la section Résolution des problèmes de Amazon VPC connexion.

Configuration d'une source Amazon Kendra de données à laquelle se connecter Amazon VPC

Lorsque vous ajoutez une nouvelle source de données Amazon Kendra, vous pouvez utiliser la Amazon VPC fonctionnalité si le connecteur de source de données sélectionné prend en charge cette fonctionnalité.

Vous pouvez configurer une nouvelle source de Amazon Kendra données Amazon VPC activée à l'aide de l'API AWS Management Console ou de l' Amazon Kendra API. Plus précisément, utilisez l'opération CreateDataSourceAPI, puis utilisez le `VpcConfiguration` paramètre pour fournir les informations suivantes :

- `SubnetIds`— Une liste d'identifiants de sous-réseaux Amazon VPC
- `SecurityGroupIds`— Une liste d'identifiants de groupes de Amazon VPC sécurité

Si vous utilisez la console, vous fournissez les Amazon VPC informations requises lors de la configuration du connecteur. Pour utiliser la console afin d'activer la fonctionnalité Amazon VPC pour un connecteur, vous devez d'abord choisir un Amazon VPC. Ensuite, vous fournissez les identifiants

de tous les sous-réseaux Amazon VPC et les identifiants de tous les groupes de sécurité Amazon VPC. Vous pouvez choisir les sous-réseaux Amazon VPC et les groupes de sécurité Amazon VPC que vous avez créés dans Configuration d'Amazon VPC, ou utiliser les sous-réseaux existants.

Rubriques

- Affichage des Amazon VPC identifiants
- Vérification de votre IAM rôle dans la source de données

Affichage des Amazon VPC identifiants

Les identifiants des sous-réseaux et des groupes de sécurité sont configurés dans la Amazon VPC console. Pour consulter les identifiants, suivez les procédures ci-dessous.

Pour afficher les identificateurs de sous-réseau

1. Connectez-vous à la console Amazon VPC AWS Management Console et ouvrez-la à l'adresse. https://console.aws.amazon.com/vpc/
2. Dans le volet de navigation, sélectionnez Subnets.
3. Dans la liste des sous-réseaux, sélectionnez le sous-réseau qui contient votre serveur de base de données.
4. Dans l'onglet Détails, notez l'identifiant dans le champ ID du sous-réseau.

Pour afficher les identifiants des groupes de sécurité

1. Connectez-vous à la console Amazon VPC AWS Management Console et ouvrez-la à l'adresse. https://console.aws.amazon.com/vpc/
2. Dans le volet de navigation, choisissez Security groups.
3. Dans la liste des groupes de sécurité, choisissez le groupe pour lequel vous souhaitez obtenir l'identifiant.
4. Dans l'onglet Détails, notez l'identifiant dans le champ ID du groupe de sécurité.

Vérification de votre IAM rôle dans la source de données

Assurez-vous que votre rôle de source de données AWS Identity and Access Management IAM(connecteur) contient les autorisations nécessaires pour accéder à votre Amazon VPC.

Si vous utilisez la console pour créer un nouveau rôle pour votre IAM rôle, ajoute Amazon Kendra automatiquement les autorisations appropriées à votre IAM rôle en votre nom. Si vous utilisez l'API ou utilisez un IAM rôle existant, vérifiez que votre rôle contient des autorisations d'accès Amazon VPC. Pour vérifier que vous disposez des autorisations appropriées, consultez la section IAM Rôles du VPC.

Vous pouvez modifier une source de données existante pour utiliser un Amazon VPC sous-réseau différent. Vérifiez toutefois le IAM rôle de votre source de données et, si nécessaire, modifiez-le pour qu'il reflète le changement afin que le connecteur de source de Amazon Kendra données fonctionne correctement.

Connexion à une base de données dans un VPC

L'exemple suivant montre comment connecter une MySQL base de données exécutée dans un cloud privé virtuel (VPC). L'exemple suppose que vous commencez par votre VPC par défaut et que vous devez créer une MySQL base de données. Si vous possédez déjà un VPC, assurez-vous qu'il est configuré comme indiqué. Si vous avez une MySQL base de données, vous pouvez l'utiliser au lieu d'en créer une nouvelle.

Étapes

- Étape 1 : Configuration d'un VPC
- Étape 2 : créer et configurer des groupes de sécurité
- Étape 3 : Création d'une base de données
- Étape 4 : Création d'un connecteur de source de données

Étape 1 : Configuration d'un VPC

Configurez votre VPC de manière à disposer d'un sous-réseau privé et d'un groupe de sécurité pour accéder Amazon Kendra à une MySQL base de données exécutée dans le sous-réseau. Les sous-réseaux fournis dans la configuration VPC doivent se trouver dans la région USA Ouest (Oregon), la région USA Est (Virginie du Nord) ou la région Europe (Irlande).

Pour configurer un VPC à l'aide de Amazon VPC

1. Connectez-vous à la console Amazon VPC AWS Management Console et ouvrez-la à l'adresse. https://console.aws.amazon.com/vpc/
2. Dans le volet de navigation, choisissez Table de routage, puis choisissez Créer une table de routage.

3. Dans le champ Nom, entrez**`Private subnet route table`**. Dans la liste déroulante VPC, sélectionnez votre VPC, puis choisissez Créer une table de routage. Choisissez Fermer pour revenir à la liste des tables de routage.
4. Dans le volet de navigation, choisissez les passerelles NAT, puis choisissez Create NAT gateway.
5. Dans la liste déroulante Sous-réseau, sélectionnez le sous-réseau public. Notez l'ID du sous-réseau.
6. Si vous n'avez pas d'adresse IP élastique, choisissez Create New EIP, Create a NAT Gateway, puis Close.
7. Dans le volet de navigation, choisissez Route tables.
8. Dans la liste des tables de routage, choisissez la table de routage du sous-réseau privé que vous avez créée à l'étape 3. Dans Actions, choisissez Modifier les itinéraires.
9. Choisissez Ajouter une route. Pour la destination, entrez **`0.0.0.0/0`** pour autoriser tout le trafic sortant vers Internet. Pour Target, choisissez NAT Gateway, puis choisissez la passerelle que vous avez créée à l'étape 4. Choisissez Enregistrer les modifications, puis Fermer.
10. Dans Actions, choisissez Modifier les associations de sous-réseaux.
11. Choisissez les sous-réseaux que vous souhaitez rendre privés. Ne choisissez pas le sous-réseau avec la passerelle NAT que vous avez mentionnée précédemment. Choisissez Enregistrer les associations lorsque vous avez terminé.

Étape 2 : créer et configurer des groupes de sécurité

Configurez ensuite les groupes de sécurité pour votre base de données.

Pour créer et configurer des groupes de sécurité

1. Connectez-vous à la console Amazon VPC AWS Management Console et ouvrez-la à l'adresse. https://console.aws.amazon.com/vpc/
2. Dans la description de votre VPC, notez le IPv4 CIDR.
3. Dans le volet de navigation, choisissez Security groups, puis Create security group.
4. Sous Security group name (Nom du groupe de sécurité), saisissez **`DataSourceInboundSecurityGroup`**. Fournissez une description, puis choisissez votre VPC dans la liste. Choisissez Créer un groupe de sécurité, puis Fermer.
5. Choisissez l’onglet Inbound rules (Règles entrantes).

6. Choisissez Modifier les règles de trafic entrant, puis sélectionnez Ajouter une règle
7. Pour une base de données, entrez le numéro de port pour la plage de ports. Par exemple, pour MySQL c'est**3306**, et pour HTTPS, c'est**443**. Pour la source, tapez le routage interdomaine sans classe (CIDR) de votre VPC. Choisissez Enregistrer les règles, puis Fermer.

Le groupe de sécurité permet à tous les membres du VPC de se connecter à la base de données et autorise les connexions sortantes vers Internet.

Étape 3 : Création d'une base de données

Créez une base de données pour stocker vos documents, ou vous pouvez utiliser votre base de données existante.

Pour obtenir des instructions sur la création d'une MySQL base de données, consultez MySQL.

Étape 4 : Création d'un connecteur de source de données

Après avoir configuré votre VPC et créé votre base de données, vous pouvez créer un connecteur de source de données pour la base de données. Pour plus d'informations sur les connecteurs de base de données compatibles Amazon Kendra , voir Connecteurs pris en charge.

Pour votre base de données, assurez-vous de configurer votre VPC, les sous-réseaux privés que vous avez créés dans votre VPC et le groupe de sécurité que vous avez créé dans votre VPC.

Résolution des problèmes de connexion VPC

Si vous rencontrez des problèmes avec votre connexion au cloud privé virtuel (VPC), vérifiez que vos IAM autorisations, les paramètres du groupe de sécurité et les tables de routage du sous-réseau sont correctement configurés.

L'une des causes potentielles de l'échec de la synchronisation du connecteur de source de données est que la source de données est peut-être inaccessible depuis le sous-réseau auquel vous l'avez affectée. Amazon Kendra Pour résoudre ce problème, nous vous recommandons de créer une Amazon EC2 instance avec les mêmes Amazon VPC paramètres. Essayez ensuite d'accéder à la source de données depuis cette Amazon EC2 instance à l'aide d'appels d'API REST ou d'autres méthodes (en fonction du type spécifique de votre source de données).

Si vous accédez à la source de données depuis l' Amazon EC2 instance que vous créez, cela signifie que votre source de données est accessible depuis ce sous-réseau. Par conséquent, votre problème

de synchronisation n'est pas lié au fait que votre source de données est inaccessible par Amazon VPC.

Si vous ne pouvez pas accéder à votre Amazon EC2 instance depuis la configuration de votre VPC et la valider avec l' Amazon EC2 instance que vous avez créée, vous devez poursuivre le dépannage. Par exemple, si la synchronisation d'un Amazon S3 connecteur a échoué en raison d'erreurs liées à des problèmes de connexion, vous pouvez configurer une Amazon EC2 instance avec la même Amazon VPC configuration que celle que vous avez attribuée à votre Amazon S3 connecteur. Utilisez ensuite cette EC2 instance Amazon pour vérifier si vous avez Amazon VPC été correctement configurée.

Voici un exemple de configuration d'une Amazon EC2 instance pour résoudre les problèmes de Amazon VPC connexion à une source de Amazon S3 données.

Rubriques

- Étape 1 : Lancer une Amazon EC2 instance
- Étape 2 : Se connecter à l' Amazon EC2 instance
- Étape 3 : Tester Amazon S3 l'accès

Étape 1 : Lancer une Amazon EC2 instance

1. Connectez-vous à la EC2 console Amazon AWS Management Console et ouvrez-la à l'adresse https://console.aws.amazon.com/ec2/.
2. Sélectionnez Lancer une instance.
3. Choisissez Paramètres réseau, puis Modifier, puis procédez comme suit :

    a. Choisissez le même VPC et le même sous-réseau que ceux auxquels vous avez été affectés. Amazon Kendra

    b. Pour Firewall (security groups) (Pare-feu (groupes de sécurité), choisissez Select existing security group (Sélectionner un groupe de sécurité existant). Sélectionnez ensuite le groupe de sécurité auquel vous avez assigné Amazon Kendra.

    > **Note**
    >
    > Le groupe de sécurité doit autoriser le trafic sortant à Amazon S3.

    c. Définissez l'attribution automatique de l'adresse IP publique sur Désactiver.

d. Dans Détails avancés, procédez comme suit :

- Pour le profil d'instance IAM, sélectionnez Créer un nouveau profil IAM pour créer et associer un profil d' IAM instance à votre instance. Assurez-vous que le profil dispose des autorisations d'accès Amazon S3. Pour plus d'informations, consultez Comment puis-je accorder à mon Amazon EC2 instance l'accès à un Amazon S3 bucket ? dans AWS re:Post.
- Conservez tous les autres paramètres par défaut.

e. Vérifiez et lancez l' Amazon EC2 instance.

Étape 2 : Se connecter à l' Amazon EC2 instance

Une fois que votre Amazon EC2 instance est en cours d'exécution, rendez-vous sur la page détaillée de votre instance et connectez-vous à votre instance. Pour ce faire, suivez les étapes décrites dans Connect to your instances sans avoir besoin d' IPv4 adresse publique à l'aide du point de terminaison EC2 Instance Connect dans le guide de Amazon EC2 l'utilisateur pour les instances Linux.

Étape 3 : Tester Amazon S3 l'accès

Une fois connecté à votre terminal d' Amazon EC2 instance, exécutez une AWS CLI commande pour tester la connexion entre ce sous-réseau privé et votre Amazon S3 bucket.

Pour tester Amazon S3 l'accès, tapez la AWS CLI commande suivante dans le AWS CLI : `aws s3 ls`

Une fois la AWS CLI commande exécutée, passez en revue les points suivants :

- Si vous avez correctement configuré les IAM autorisations nécessaires et que votre Amazon S3 configuration est correcte, vous devriez voir la liste de vos Amazon S3 buckets.
- Si vous constatez des erreurs d'autorisation, par exemple`Access Denied`, il est probable que la configuration de votre VPC soit correcte, mais qu'il y a un problème avec vos IAM autorisations ou votre politique de Amazon S3 compartiment.

Si le délai de la commande est expiré, il est probable que votre connexion soit interrompue car la configuration de votre VPC est incorrecte et l' EC2 instance Amazon ne peut pas accéder à Amazon S3 depuis votre sous-réseau. Reconfigurez votre VPC, puis réessayez.

Supprimer un index, une source de données ou des documents téléchargés par lots

Cette section explique comment supprimer un index, un référentiel de sources de données contenant des documents de votre index ou des documents de votre index que vous avez chargés par lots.

Rubriques

- Supprimer un index
- Suppression d'une source de données
- Suppression de documents téléchargés par lots

Supprimer un index

Vous pouvez supprimer un index Amazon Kendra lorsque vous ne l'utilisez plus. Par exemple, supprimez un index lorsque :

- Vous n'utilisez plus l'indice et souhaitez réduire les frais de votre AWS compte. Un Amazon Kendra index génère des frais pendant son exécution, que vous fassiez ou non des requêtes sur l'index.
- Vous souhaitez reconfigurer l'index pour une autre édition de Amazon Kendra. Supprimez l'index existant, puis créez-en un nouveau avec l'édition différente.
- Vous avez atteint le nombre maximum d'index dans votre compte et vous ne souhaitez pas dépasser votre quota. Supprimez un index existant et ajoutez-en un nouveau. Pour plus d'informations sur le nombre maximal d'index que vous pouvez créer, consultez la section Quotas.

Pour supprimer un index, utilisez la console AWS Command Line Interface, le AWS CloudFormation script ou l'`DeleteIndex`API. La suppression d'un index entraîne la suppression de l'index ainsi que de toutes les sources de données et données de document associées. La suppression d'un index ne supprime pas les documents originaux de votre espace de stockage.

La suppression d'un index est une opération asynchrone. Lorsque vous commencez à supprimer un index, le statut de l'index passe à`DELETING`. Il reste dans `DELETING` cet état jusqu'à ce que toutes les informations relatives à l'index soient supprimées. Une fois l'index supprimé, il n'apparaît plus dans les résultats d'un appel à l'ListIndicesAPI. Si vous appelez l'DescribeIndexAPI avec l'identifiant de l'index supprimé, vous recevez `ResourceNotFound` une exception.

Pour supprimer un index (console)

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la à l'adresse https://console.aws.amazon.com/kendra/.
2. Dans le volet de navigation, choisissez Indexes, puis choisissez l'index à supprimer.
3. Choisissez Supprimer pour supprimer l'index sélectionné.

Pour supprimer un index (CLI)

- Dans le AWS CLI, utilisez la commande suivante. La commande est formatée pour Linux et macOS. Si vous utilisez Windows, remplacez le caractère de continuation de ligne Unix (\) par un curseur (^).

```
aws kendra delete-index \
   --id index-id
```

Suppression d'une source de données

Vous supprimez une source de données lorsque vous souhaitez supprimer les informations qu'elle contient de votre Amazon Kendra index. Supprimez par exemple une source de données lorsque :

- Une source de données n'est pas correctement configurée. Supprimez la source de données, attendez que sa suppression soit terminée, puis recréez-la.
- Vous avez migré des documents d'une source de données à une autre. Supprimez la source de données d'origine et recréez-la dans le nouvel emplacement.
- Vous avez atteint le nombre limite de sources de données pour un index. Supprimez l'une des sources de données existantes et ajoutez-en une nouvelle. Pour plus d'informations sur le nombre de sources de données que vous pouvez créer, consultezQuotas.

Pour supprimer une source de données, utilisez la console, le AWS Command Line Interface (AWS CLI), l'`DeleteDataSource`API ou un AWS CloudFormation script. La suppression d'une source de données supprime toutes les informations relatives à la source de données de l'index. Si vous souhaitez uniquement arrêter la synchronisation de la source de données, modifiez le calendrier de synchronisation de la source de données pour « exécuter à la demande ».

La suppression d'une source de données est une opération asynchrone. Lorsque vous commencez à supprimer une source de données, son statut passe à`DELETING`. Il reste dans `DELETING` cet état jusqu'à ce que les informations relatives à la source de données soient supprimées. Une fois la source de données supprimée, elle n'apparaît plus dans les résultats d'un appel à l'ListDataSourcesAPI. Si vous appelez l'DescribeDataSourceAPI avec l'identifiant de la source de données supprimée, vous recevez une `ResourceNotFound` exception.

> **Note**
>
> La suppression d'une source de données complète ou la resynchronisation de votre index après avoir supprimé des documents spécifiques d'une source de données peut prendre jusqu'à une heure, voire plus, selon le nombre de documents que vous souhaitez supprimer.

Pour supprimer une source de données (console)

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la à l'adresse https://console.aws.amazon.com/kendra/.
2. Dans le volet de navigation, choisissez Indexes, puis choisissez l'index contenant la source de données à supprimer.
3. Dans le panneau de navigation, choisissez Sources de données.
4. Choisissez la source de données à supprimer.
5. Choisissez Supprimer pour supprimer la source de données.

Pour supprimer une source de données (CLI)

- Dans le AWS Command Line Interface, utilisez la commande suivante. La commande est formatée pour Linux et macOS. Si vous utilisez Windows, remplacez le caractère de continuation de ligne Unix (\) par un curseur (^).

```
aws kendra delete-data-source \
   --id data-source-id \
   --index-id index-id
```

Lorsque vous supprimez une source de données, Amazon Kendra toutes les informations stockées la concernant sont supprimées. Amazon Kendra supprime toutes les données du document stockées

dans l'index, ainsi que tous les historiques d'exécution et les mesures associés à la source de données. La suppression d'une source de données ne supprime pas les documents originaux de votre espace de stockage.

Les documents de la source de données peuvent être inclus dans le nombre de documents renvoyé par l'`DescribeIndex`API lors de la Amazon Kendra suppression d'une source de données. Les documents de la source de données peuvent apparaître dans les résultats de recherche lorsque la source Amazon Kendra de données est supprimée.

Amazon Kendra libère les ressources d'une source de données dès que vous appelez l'`DeleteDataSource`API ou que vous choisissez de supprimer la source de données dans la console. Si vous supprimez la source de données pour réduire le nombre de sources de données en dessous de votre limite, vous pouvez créer une nouvelle source de données immédiatement.

Si vous supprimez une source de données puis que vous créez une autre source de données pour les données du document, attendez que la première source de données soit supprimée avant de synchroniser la nouvelle source de données.

Vous pouvez supprimer une source de données en cours de synchronisation. Amazon Kendra La synchronisation est arrêtée et la source de données est supprimée. Si vous essayez de démarrer une synchronisation alors que la source de données est supprimée, vous obtenez une `ConflictException` exception.

Vous ne pouvez pas supprimer une source de données si l'index associé est dans `DELETING` cet état. La suppression d'un index entraîne la suppression de toutes les sources de données de l'index. Vous pouvez commencer à supprimer un index alors qu'une source de données correspondant à cet index est dans son `DELETING` état actuel.

Si deux sources de données pointent vers les mêmes documents, par exemple deux sources de données pointant vers le même Amazon S3 compartiment, les documents de l'index peuvent être incohérents lorsque l'une des sources de données est supprimée. Lorsque deux sources de données font référence aux mêmes documents, une seule copie des données du document est stockée dans l'index. La suppression d'une source de données entraîne la suppression des données d'index des documents. L'autre source de données ne sait pas que les documents ont été supprimés et Amazon Kendra ne les réindexera donc pas correctement lors de la prochaine synchronisation. Lorsque deux sources de données pointent vers le même emplacement de document, vous devez supprimer les deux sources de données, puis en recréer une.

Suppression de documents téléchargés par lots

Vous pouvez supprimer des documents directement depuis un index à l'aide de l'BatchDeleteDocumentAPI. Vous ne pouvez pas supprimer de documents directement à l'aide de la console. Si vous utilisez la console, vous pouvez soit supprimer des documents spécifiques de votre référentiel de sources de données et les resynchroniser avec votre index, soit supprimer l'intégralité du connecteur de source de données.

La suppression de documents d'un index à l'aide d'une opération asynchrone `BatchDeleteDocument` est une opération asynchrone. Après avoir appelé l'`BatchDeleteDocument`API, vous l'BatchGetDocumentStatusutilisez pour suivre la progression de la suppression de vos documents. Lorsqu'un document est supprimé de l'index, Amazon Kendra renvoie `NOT_FOUND` son statut.

> **Note**
>
> La suppression de documents d'un index `BatchDeleteDocument` peut prendre jusqu'à une heure ou plus, selon le nombre de documents que vous souhaitez supprimer.

Pour supprimer les documents téléchargés par lots d'un index (CLI)

- Dans le AWS Command Line Interface, utilisez la commande suivante. La commande est formatée pour Linux et macOS. Si vous utilisez Windows, remplacez le caractère de continuation de ligne Unix (\) par un curseur (^).

```
aws kendra batch-delete-document \
   --index-id index-id \
   --document-id-list 'doc-id-1' 'doc-id-2'
```

Enrichir vos documents lors de l'ingestion

> **Note**
>
> La prise en charge des fonctionnalités varie en fonction du type d'index et de l'API de recherche utilisés. Pour savoir si cette fonctionnalité est prise en charge pour le type d'index et l'API de recherche que vous utilisez, consultez la section Types d'index.

Vous pouvez modifier le contenu et les champs ou attributs des métadonnées du document pendant le processus d'ingestion du document. Grâce à Amazon Kendra la fonction d'enrichissement personnalisé des documents, vous pouvez créer, modifier ou supprimer les attributs et le contenu des documents lorsque vous les ingérez. Amazon Kendra Cela signifie que vous pouvez manipuler et ingérer vos données selon vos besoins.

Cette fonctionnalité vous permet de contrôler la manière dont vos documents sont traités et ingérés. Amazon Kendra Par exemple, vous pouvez supprimer les informations personnellement identifiables contenues dans les métadonnées du document lors de l'ingestion de vos documents. Amazon Kendra

Vous pouvez également utiliser cette fonctionnalité en invoquant une fonction Lambda AWS Lambda pour exécuter la reconnaissance optique de caractères (OCR) sur des images, une traduction sur du texte et d'autres tâches de préparation des données pour la recherche ou l'analyse. Par exemple, vous pouvez appeler une fonction pour exécuter l'OCR sur des images. La fonction pouvait interpréter le texte des images et traiter chaque image comme un document textuel. Une entreprise qui reçoit des enquêtes clients envoyées par la poste et qui les stocke sous forme d'images pourrait intégrer ces images sous forme de documents textuels. Amazon Kendra L'entreprise peut ensuite rechercher des informations précieuses issues de sondages auprès des clients dans Amazon Kendra.

Vous pouvez utiliser des opérations de base à appliquer lors de la première analyse de vos données, puis utiliser une fonction Lambda pour appliquer des opérations plus complexes à vos données. Par exemple, vous pouvez utiliser une opération de base pour simplement supprimer toutes les valeurs du champ de métadonnées du document « Customer_ID », puis appliquer une fonction Lambda pour extraire le texte des images du texte des documents.

Comment fonctionne l'enrichissement personnalisé des documents

Le processus global d'enrichissement de documents personnalisés est le suivant :

1. Vous configurez l'enrichissement personnalisé des documents lorsque vous créez ou mettez à jour votre source de données, ou lorsque vous y indexez directement vos documents Amazon Kendra.
2. Amazon Kendra applique des configurations intégrées ou une logique de base pour modifier vos données. Pour de plus amples informations, veuillez consulter the section called “Opérations de base pour modifier les métadonnées”.
3. Si vous choisissez de configurer la manipulation avancée des données, vous Amazon Kendra pouvez l'appliquer à vos documents bruts originaux ou aux documents structurés et analysés. Pour de plus amples informations, veuillez consulter the section called “Fonctions Lambda : extraire et modifier les métadonnées ou le contenu”.
4. Vos documents modifiés sont ingérés dans. Amazon Kendra

À tout moment de ce processus, si votre configuration n'est pas valide, une Amazon Kendra erreur est générée.

Lorsque vous appelez CreateDataSource, ou UpdateDataSourceBatchPutDocument APIs, vous fournissez votre configuration personnalisée d'enrichissement de documents. Si vous appelez`BatchPutDocument`, vous devez configurer l'enrichissement personnalisé des documents à chaque demande. Si vous utilisez la console, vous sélectionnez votre index, puis sélectionnez Enrichissements de documents pour configurer l'enrichissement de documents personnalisé.

Si vous utilisez les enrichissements de documents dans la console, vous pouvez choisir de configurer uniquement les opérations de base ou uniquement les fonctions Lambda, ou les deux, comme vous pouvez le faire avec l'API. Vous pouvez sélectionner Suivant dans les étapes de la console pour choisir de ne pas configurer les opérations de base et uniquement les fonctions Lambda, notamment si elles doivent s'appliquer aux données d'origine (pré-extraction) ou structurées (après extraction). Vous ne pouvez enregistrer vos configurations qu'en effectuant toutes les étapes dans la console. Les configurations de vos documents ne sont pas enregistrées si vous n'effectuez pas toutes les étapes.

Opérations de base pour modifier les métadonnées

Vous pouvez manipuler les champs et le contenu de votre document à l'aide d'une logique de base. Cela inclut la suppression de valeurs dans un champ, la modification des valeurs d'un champ à l'aide

d'une condition ou la création d'un champ. Pour les manipulations avancées qui vont au-delà de ce que vous pouvez manipuler en utilisant la logique de base, appelez une fonction Lambda. Pour de plus amples informations, veuillez consulter the section called “Fonctions Lambda : extraire et modifier les métadonnées ou le contenu”.

Pour appliquer la logique de base, vous devez spécifier le champ cible que vous souhaitez manipuler à l'aide de l'DocumentAttributeTargetobjet. Vous fournissez la clé d'attribut. Par exemple, la clé « Département » est un champ ou un attribut qui contient tous les noms de départements associés aux documents. Vous pouvez également spécifier une valeur à utiliser dans le champ cible si une certaine condition est remplie. Vous définissez la condition à l'aide de l'DocumentAttributeConditionobjet. Par exemple, si le champ « Source_URI » contient « financier » dans sa valeur d'URI, préremplissez le champ cible « Department » avec la valeur cible « Finance » pour le document. Vous pouvez également supprimer les valeurs de l'attribut du document cible.

Pour appliquer une logique de base à l'aide de la console, sélectionnez votre index, puis sélectionnez Enrichissements de documents dans le menu de navigation. Accédez à Configurer les opérations de base pour appliquer des manipulations de base aux champs et au contenu de votre document.

Voici un exemple d'utilisation de la logique de base pour supprimer tous les numéros d'identification des clients dans le champ du document appelé « Customer_ID ».

Exemple 1 : Supprimer les numéros d'identification des clients associés aux documents

Données avant l'application de la manipulation de base.

Identifiant du document	Corps_Text	Identifiant_client
1	Lorem Ipsum.	CID1234
2	Lorem Ipsum.	CID1235
3	Lorem Ipsum.	CID1236

Données appliquées après manipulation de base.

Identifiant du document	Corps_Text	Identifiant_client
1	Lorem Ipsum.	

Identifiant du document	Corps_Text	Identifiant_client
2	Lorem Ipsum.	
3	Lorem Ipsum.	

Voici un exemple d'utilisation de la logique de base pour créer un champ appelé « Department » et préremplir ce champ avec les noms des départements en fonction des informations du champ « Source_URI ». Cela utilise la condition selon laquelle si le champ « Source_URI » contient « financier » dans sa valeur d'URI, le champ cible « Department » est prérempli avec la valeur cible « Finance » pour le document.

Exemple 2 : créer le champ « Département » et le préremplir avec les noms des départements associés aux documents à l'aide d'une condition.

Données avant l'application de la manipulation de base.

Identifiant du document	Corps_Text	URI de la source
1	Lorem Ipsum.	financier/1
2	Lorem Ipsum.	financier/2
3	Lorem Ipsum.	financier/3

Données appliquées après manipulation de base.

Identifiant du document	Corps_Text	URI de la source	Département
1	Lorem Ipsum.	financier/1	Finance
2	Lorem Ipsum.	financier/2	Finance
3	Lorem Ipsum.	financier/3	Finance

> **Note**
>
> Amazon Kendra Impossible de créer un champ de document cible s'il n'est pas déjà créé en tant que champ d'index. Après avoir créé votre champ d'index, vous pouvez créer un champ de document à l'aide de`DocumentAttributeTarget`. Amazon Kendra fait ensuite correspondre le champ de métadonnées du document que vous venez de créer à votre champ d'index.

Le code suivant est un exemple de configuration de la manipulation de base des données pour supprimer les numéros d'identification des clients associés aux documents.

Console

Pour configurer la manipulation de base des données afin de supprimer les numéros d'identification des clients

1. Dans le volet de navigation de gauche, sous Index, sélectionnez Enrichissements de documents, puis sélectionnez Ajouter un enrichissement de document.
2. Sur la page Configurer les opérations de base, choisissez dans le menu déroulant la source de données dont vous souhaitez modifier les champs et le contenu du document. Choisissez ensuite dans le menu déroulant le nom du champ du document « Customer_ID », sélectionnez dans le menu déroulant le nom du champ d'index « Customer_ID » et sélectionnez dans le menu déroulant l'action cible Supprimer. Sélectionnez ensuite Ajouter une opération de base.

CLI

Pour configurer la manipulation de base des données afin de supprimer les numéros d'identification des clients

```
aws kendra create-data-source \
 --name data-source-name \
 --index-id index-id \
 --role-arn arn:aws:iam::account-id:role/role-name \
 --type S3 \
 --configuration '{"S3Configuration":{"BucketName":"S3-bucket-name"}}' \
```

```
 --custom-document-enrichment-configuration '{"InlineConfigurations":[{"Target":
{"TargetDocumentAttributeKey":"Customer_ID", "TargetDocumentAttributeValueDeletion":
 true}}]}'
```

Python

Pour configurer la manipulation de base des données afin de supprimer les numéros d'identification des clients

```
import boto3
from botocore.exceptions import ClientError
import pprint
import time

kendra = boto3.client("kendra")

print("Create a data source with customizations")

# Provide the name of the data source
name = "data-source-name"
# Provide the index ID for the data source
index_id = "index-id"
# Provide the IAM role ARN required for data sources
role_arn = "arn:aws:iam::${account-id}:role/${role-name}"
# Provide the data source connection information
data_source_type = "S3"
S3_bucket_name = "S3-bucket-name"
# Configure the data source with Custom Document Enrichment
configuration = {"S3Configuration":
        {
            "BucketName": S3_bucket_name
        }
    }
custom_document_enrichment_configuration = {"InlineConfigurations":[
        {
            "Target":{"TargetDocumentAttributeKey":"Customer_ID",
                      "TargetDocumentAttributeValueDeletion": True}
        }]
    }

try:
    data_source_response = kendra.create_data_source(
        Name = name,
```

```
        IndexId = index_id,
        RoleArn = role_arn,
        Type = data_source_type
        Configuration = configuration
        CustomDocumentEnrichmentConfiguration =
custom_document_enrichment_configuration
    )

    pprint.pprint(data_source_response)

    data_source_id = data_source_response["Id"]

    print("Wait for Amazon Kendra to create the data source with your
customizations.")

    while True:
        # Get the details of the data source, such as the status
        data_source_description = kendra.describe_data_source(
            Id = data_source_id,
            IndexId = index_id
        )
        status = data_source_description["Status"]
        print(" Creating data source. Status: "+status)
        time.sleep(60)
        if status != "CREATING":
            break

    print("Synchronize the data source.")

    sync_response = kendra.start_data_source_sync_job(
        Id = data_source_id,
        IndexId = index_id
    )

    pprint.pprint(sync_response)

    print("Wait for the data source to sync with the index.")

    while True:

        jobs = kendra.list_data_source_sync_jobs(
            Id= data_source_id,
            IndexId= index_id
        )
```

```
        # For this example, there should be one job
        status = jobs["History"][0]["Status"]

        print(" Syncing data source. Status: "+status)
        time.sleep(60)
        if status != "SYNCING":
            break

except  ClientError as e:
        print("%s" % e)

print("Program ends.")
```

Java

Pour configurer la manipulation de base des données afin de supprimer les numéros d'identification des clients

```
package com.amazonaws.kendra;

import java.util.concurrent.TimeUnit;
import software.amazon.awssdk.services.kendra.KendraClient;
import software.amazon.awssdk.services.kendra.model.CreateDataSourceRequest;
import software.amazon.awssdk.services.kendra.model.CreateDataSourceResponse;
import software.amazon.awssdk.services.kendra.model.CreateIndexRequest;
import software.amazon.awssdk.services.kendra.model.CreateIndexResponse;
import software.amazon.awssdk.services.kendra.model.DataSourceConfiguration;
import software.amazon.awssdk.services.kendra.model.DataSourceStatus;
import software.amazon.awssdk.services.kendra.model.DataSourceSyncJob;
import software.amazon.awssdk.services.kendra.model.DataSourceSyncJobStatus;
import software.amazon.awssdk.services.kendra.model.DataSourceType;
import software.amazon.awssdk.services.kendra.model.DescribeDataSourceRequest;
import software.amazon.awssdk.services.kendra.model.DescribeDataSourceResponse;
import software.amazon.awssdk.services.kendra.model.DescribeIndexRequest;
import software.amazon.awssdk.services.kendra.model.DescribeIndexResponse;
import software.amazon.awssdk.services.kendra.model.IndexStatus;
import software.amazon.awssdk.services.kendra.model.ListDataSourceSyncJobsRequest;
import software.amazon.awssdk.services.kendra.model.ListDataSourceSyncJobsResponse;
import software.amazon.awssdk.services.kendra.model.S3DataSourceConfiguration;
import software.amazon.awssdk.services.kendra.model.StartDataSourceSyncJobRequest;
import software.amazon.awssdk.services.kendra.model.StartDataSourceSyncJobResponse;
```

```
public class CreateDataSourceWithCustomizationsExample {

    public static void main(String[] args) throws InterruptedException {
        System.out.println("Create a data source with customizations");

        String dataSourceName = "data-source-name";
        String indexId = "index-id";
        String dataSourceRoleArn = "arn:aws:iam::account-id:role/role-name";
        String s3BucketName = "S3-bucket-name"

        KendraClient kendra = KendraClient.builder().build();

        CreateDataSourceRequest createDataSourceRequest = CreateDataSourceRequest
            .builder()
            .name(dataSourceName)
            .description(experienceDescription)
            .roleArn(experienceRoleArn)
            .type(DataSourceType.S3)
            .configuration(
                DataSourceConfiguration
                    .builder()
                    .s3Configuration(
                        S3DataSourceConfiguration
                            .builder()
                            .bucketName(s3BucketName)
                            .build()
                    ).build()
            )
            .customDocumentEnrichmentConfiguration(
                CustomDocumentEnrichmentConfiguration
                    .builder()
                    .inlineConfigurations(Arrays.asList(
                        InlineCustomDocumentEnrichmentConfiguration
                            .builder()
                            .target(
                                DocumentAttributeTarget
                                    .builder()
                                    .targetDocumentAttributeKey("Customer_ID")
                                    .targetDocumentAttributeValueDeletion(true)
                                    .build())
                            .build()
                    )).build();
```

```
        CreateDataSourceResponse createDataSourceResponse =
kendra.createDataSource(createDataSourceRequest);
        System.out.println(String.format("Response of creating data source: %s",
createDataSourceResponse));

        String dataSourceId = createDataSourceResponse.id();
        System.out.println(String.format("Waiting for Kendra to create the data
source %s", dataSourceId));
        DescribeDataSourceRequest describeDataSourceRequest =
DescribeDataSourceRequest
            .builder()
            .indexId(indexId)
            .id(dataSourceId)
            .build();

        while (true) {
            DescribeDataSourceResponse describeDataSourceResponse =
kendra.describeDataSource(describeDataSourceRequest);

            DataSourceStatus status = describeDataSourceResponse.status();
            System.out.println(String.format("Creating data source. Status: %s",
status));
            TimeUnit.SECONDS.sleep(60);
            if (status != DataSourceStatus.CREATING) {
                break;
            }
        }

        System.out.println(String.format("Synchronize the data source %s",
dataSourceId));
        StartDataSourceSyncJobRequest startDataSourceSyncJobRequest =
StartDataSourceSyncJobRequest
            .builder()
            .indexId(indexId)
            .id(dataSourceId)
            .build();
        StartDataSourceSyncJobResponse startDataSourceSyncJobResponse =
kendra.startDataSourceSyncJob(startDataSourceSyncJobRequest);
        System.out.println(String.format("Waiting for the data
source to sync with the index %s for execution ID %s", indexId,
startDataSourceSyncJobResponse.executionId()));

        // For this example, there should be one job
```

```
        ListDataSourceSyncJobsRequest listDataSourceSyncJobsRequest =
 ListDataSourceSyncJobsRequest
            .builder()
            .indexId(indexId)
            .id(dataSourceId)
            .build();

        while (true) {
            ListDataSourceSyncJobsResponse listDataSourceSyncJobsResponse =
 kendra.listDataSourceSyncJobs(listDataSourceSyncJobsRequest);
            DataSourceSyncJob job = listDataSourceSyncJobsResponse.history().get(0);
            System.out.println(String.format("Syncing data source. Status: %s",
 job.status()));

            TimeUnit.SECONDS.sleep(60);
            if (job.status() != DataSourceSyncJobStatus.SYNCING) {
                break;
            }

        }

        System.out.println("Data source creation with customizations is complete");
    }
}
```

Fonctions Lambda : extraire et modifier les métadonnées ou le contenu

Vous pouvez manipuler les champs et le contenu de votre document à l'aide des fonctions Lambda. Cela est utile si vous souhaitez aller au-delà de la logique de base et appliquer des manipulations de données avancées. Par exemple, utiliser la reconnaissance optique de caractères (OCR), qui interprète le texte des images et traite chaque image comme un document textuel. Vous pouvez également récupérer la date-heure actuelle dans un certain fuseau horaire et insérer la date-heure là où il y a une valeur vide pour un champ de date.

Vous pouvez d'abord appliquer une logique de base, puis utiliser une fonction Lambda pour continuer à manipuler vos données, ou vice versa. Vous pouvez également choisir de n'appliquer qu'une fonction Lambda.

Amazon Kendra peut invoquer une fonction Lambda pour appliquer des manipulations de données avancées pendant le processus d'ingestion dans le cadre de votre. CustomDocumentEnrichmentConfiguration Vous spécifiez un rôle qui inclut l'autorisation d'exécuter la fonction Lambda et d'accéder à votre Amazon S3 bucket pour stocker le résultat de vos manipulations IAM de données (voir rôles d'accès).

Amazon Kendra peut appliquer une fonction Lambda sur vos documents bruts originaux ou sur les documents structurés et analysés. Vous pouvez configurer une fonction Lambda qui prend vos données d'origine ou brutes et applique vos manipulations de données à l'aide de. PreExtractionHookConfiguration Vous pouvez également configurer une fonction Lambda qui prend vos documents structurés et applique vos manipulations de données à l'aide de. PostExtractionHookConfiguration Amazon Kendra extrait les métadonnées et le texte du document pour structurer vos documents. Vos fonctions Lambda doivent respecter les structures de demande et de réponse obligatoires. Pour de plus amples informations, veuillez consulter the section called “Contrats de données pour les fonctions Lambda”.

Pour configurer une fonction Lambda dans la console, sélectionnez votre index, puis sélectionnez Enrichissements de documents dans le menu de navigation. Accédez à Configurer les fonctions Lambda pour configurer une fonction Lambda.

Vous ne pouvez configurer qu'une seule fonction Lambda pour `PreExtractionHookConfiguration` et une seule fonction Lambda pour. `PostExtractionHookConfiguration` Toutefois, votre fonction Lambda peut invoquer d'autres fonctions dont elle a besoin. Vous pouvez configurer les deux `PostExtractionHookConfiguration` ou `PreExtractionHookConfiguration` l'un ou l'autre. Votre fonction Lambda pour ne `PreExtractionHookConfiguration` doit pas dépasser une durée d'exécution de 5 minutes et votre fonction Lambda pour ne `PostExtractionHookConfiguration` doit pas dépasser une durée d'exécution de 1 minute. La configuration de l'enrichissement personnalisé des documents prend naturellement plus de temps pour intégrer vos documents Amazon Kendra que si vous ne le configuriez pas.

Vous pouvez configurer Amazon Kendra pour appeler une fonction Lambda uniquement si une condition est remplie. Par exemple, vous pouvez spécifier une condition selon laquelle, s'il existe des valeurs date-heure vides, invoque une fonction qui insère la date-heure actuelle. Amazon Kendra

Voici un exemple d'utilisation d'une fonction Lambda pour exécuter l'OCR afin d'interpréter du texte à partir d'images et de stocker ce texte dans un champ appelé « Document_Image_Text ».

Exemple 1 : extraction de texte à partir d'images pour créer des documents textuels

Données avant application de la manipulation avancée.

Identifiant du document	Document_Image
1	image_1.png
2	image_2.png
3	image_3.png

Données appliquées après manipulation avancée.

Identifiant du document	Document_Image	Document_Image_Texte
1	image_1.png	Réponse au sondage envoyée par la poste
2	image_2.png	Réponse au sondage envoyée par la poste
3	image_3.png	Réponse au sondage envoyée par la poste

Voici un exemple d'utilisation d'une fonction Lambda pour insérer la date-heure actuelle pour les valeurs de date vides. Cela utilise la condition selon laquelle si la valeur d'un champ de date est « nulle », remplacez-la par la date-heure actuelle.

Exemple 2 : remplacement des valeurs vides du champ Last_Updated par la date-heure actuelle.

Données avant application de la manipulation avancée.

Identifiant du document	Corps_Text	Dernière mise à jour
1	Lorem Ipsum.	1er janvier 2020
2	Lorem Ipsum.	
3	Lorem Ipsum.	1er juillet 2020

Données appliquées après manipulation avancée.

Identifiant du document	Corps_Text	Dernière mise à jour
1	Lorem Ipsum.	1er janvier 2020
2	Lorem Ipsum.	1er décembre 2021
3	Lorem Ipsum.	1er juillet 2020

Le code suivant est un exemple de configuration d'une fonction Lambda pour la manipulation avancée des données d'origine brutes.

Console

Pour configurer une fonction Lambda pour une manipulation avancée des données d'origine brutes

1. Dans le volet de navigation de gauche, sous Index, sélectionnez Enrichissements de documents, puis sélectionnez Ajouter un enrichissement de document.
2. Sur la page Configurer les fonctions Lambda, dans la section Lambda pour la pré-extraction, sélectionnez dans les listes déroulantes l'ARN de votre fonction Lambda et votre bucket. Amazon S3 Ajoutez votre rôle IAM d'accès en sélectionnant l'option permettant de créer un nouveau rôle dans le menu déroulant. Cela crée les Amazon Kendra autorisations requises pour créer l'enrichissement du document.

CLI

Pour configurer une fonction Lambda pour une manipulation avancée des données d'origine brutes

```
aws kendra create-data-source \
 --name data-source-name \
 --index-id index-id \
 --role-arn arn:aws:iam::account-id:role/role-name \
 --type S3 \
 --configuration '{"S3Configuration":{"BucketName":"S3-bucket-name"}}' \
```

```
 --custom-document-enrichment-configuration '{"PreExtractionHookConfiguration":
{"LambdaArn":"arn:aws:iam::account-id:function/function-name", "S3Bucket":"S3-
bucket-name"}, "RoleArn": "arn:aws:iam:account-id:role/cde-role-name"}'
```

Python

Pour configurer une fonction Lambda pour une manipulation avancée des données d'origine brutes

```
import boto3
from botocore.exceptions import ClientError
import pprint
import time

kendra = boto3.client("kendra")

print("Create a data source with customizations.")

# Provide the name of the data source
name = "data-source-name"
# Provide the index ID for the data source
index_id = "index-id"
# Provide the IAM role ARN required for data sources
role_arn = "arn:aws:iam::${account-id}:role/${role-name}"
# Provide the data source connection information
data_source_type = "S3"
S3_bucket_name = "S3-bucket-name"
# Configure the data source with Custom Document Enrichment
configuration = {"S3Configuration":
        {
            "BucketName": S3_bucket_name
        }
    }
custom_document_enrichment_configuration = {"PreExtractionHookConfiguration":
        {
            "LambdaArn":"arn:aws:iam::account-id:function/function-name",
            "S3Bucket":"S3-bucket-name"
        }
    "RoleArn":"arn:aws:iam::account-id:role/cde-role-name"
    }

try:
    data_source_response = kendra.create_data_source(
```

```
        Name = name,
        IndexId = index_id,
        RoleArn = role_arn,
        Type = data_source_type
        Configuration = configuration
        CustomDocumentEnrichmentConfiguration =
custom_document_enrichment_configuration
    )

    pprint.pprint(data_source_response)

    data_source_id = data_source_response["Id"]

    print("Wait for Amazon Kendra to create the data source with your
customizations.")

    while True:
        # Get the details of the data source, such as the status
        data_source_description = kendra.describe_data_source(
            Id = data_source_id,
            IndexId = index_id
        )
        status = data_source_description["Status"]
        print(" Creating data source. Status: "+status)
        time.sleep(60)
        if status != "CREATING":
            break

    print("Synchronize the data source.")

    sync_response = kendra.start_data_source_sync_job(
        Id = data_source_id,
        IndexId = index_id
    )

    pprint.pprint(sync_response)

    print("Wait for the data source to sync with the index.")

    while True:

        jobs = kendra.list_data_source_sync_jobs(
            Id = data_source_id,
            IndexId = index_id
```

```
        )

        # For this example, there should be one job
        status = jobs["History"][0]["Status"]

        print(" Syncing data source. Status: "+status)
        time.sleep(60)
        if status != "SYNCING":
            break

except  ClientError as e:
        print("%s" % e)

print("Program ends.")
```

Java

Pour configurer une fonction Lambda pour une manipulation avancée des données d'origine brutes

```
package com.amazonaws.kendra;

import java.util.concurrent.TimeUnit;
import software.amazon.awssdk.services.kendra.KendraClient;
import software.amazon.awssdk.services.kendra.model.CreateDataSourceRequest;
import software.amazon.awssdk.services.kendra.model.CreateDataSourceResponse;
import software.amazon.awssdk.services.kendra.model.CreateIndexRequest;
import software.amazon.awssdk.services.kendra.model.CreateIndexResponse;
import software.amazon.awssdk.services.kendra.model.DataSourceConfiguration;
import software.amazon.awssdk.services.kendra.model.DataSourceStatus;
import software.amazon.awssdk.services.kendra.model.DataSourceSyncJob;
import software.amazon.awssdk.services.kendra.model.DataSourceSyncJobStatus;
import software.amazon.awssdk.services.kendra.model.DataSourceType;
import software.amazon.awssdk.services.kendra.model.DescribeDataSourceRequest;
import software.amazon.awssdk.services.kendra.model.DescribeDataSourceResponse;
import software.amazon.awssdk.services.kendra.model.DescribeIndexRequest;
import software.amazon.awssdk.services.kendra.model.DescribeIndexResponse;
import software.amazon.awssdk.services.kendra.model.IndexStatus;
import software.amazon.awssdk.services.kendra.model.ListDataSourceSyncJobsRequest;
import software.amazon.awssdk.services.kendra.model.ListDataSourceSyncJobsResponse;
import software.amazon.awssdk.services.kendra.model.S3DataSourceConfiguration;
import software.amazon.awssdk.services.kendra.model.StartDataSourceSyncJobRequest;
import software.amazon.awssdk.services.kendra.model.StartDataSourceSyncJobResponse;
```

```
public class CreateDataSourceWithCustomizationsExample {

    public static void main(String[] args) throws InterruptedException {
        System.out.println("Create a data source with customizations");

        String dataSourceName = "data-source-name";
        String indexId = "index-id";
        String dataSourceRoleArn = "arn:aws:iam::account-id:role/role-name";
        String s3BucketName = "S3-bucket-name"

        KendraClient kendra = KendraClient.builder().build();

        CreateDataSourceRequest createDataSourceRequest = CreateDataSourceRequest
            .builder()
            .name(dataSourceName)
            .description(experienceDescription)
            .roleArn(experienceRoleArn)
            .type(DataSourceType.S3)
            .configuration(
                DataSourceConfiguration
                    .builder()
                    .s3Configuration(
                        S3DataSourceConfiguration
                            .builder()
                            .bucketName(s3BucketName)
                            .build()
                    ).build()
            )
            .customDocumentEnrichmentConfiguration(
                CustomDocumentEnrichmentConfiguration
                    .builder()
                    .preExtractionHookConfiguration(
                        HookConfiguration
                            .builder()
                            .lambdaArn("arn:aws:iam::account-id:function/function-
name")
                            .s3Bucket("S3-bucket-name")
                            .build())
                    .roleArn("arn:aws:iam::account-id:role/cde-role-name")
                    .build();
```

```
        CreateDataSourceResponse createDataSourceResponse =
kendra.createDataSource(createDataSourceRequest);
        System.out.println(String.format("Response of creating data source: %s",
createDataSourceResponse));

        String dataSourceId = createDataSourceResponse.id();
        System.out.println(String.format("Waiting for Kendra to create the data
source %s", dataSourceId));
        DescribeDataSourceRequest describeDataSourceRequest =
DescribeDataSourceRequest
            .builder()
            .indexId(indexId)
            .id(dataSourceId)
            .build();

        while (true) {
            DescribeDataSourceResponse describeDataSourceResponse =
kendra.describeDataSource(describeDataSourceRequest);

            DataSourceStatus status = describeDataSourceResponse.status();
            System.out.println(String.format("Creating data source. Status: %s",
status));
            TimeUnit.SECONDS.sleep(60);
            if (status != DataSourceStatus.CREATING) {
                break;
            }
        }

        System.out.println(String.format("Synchronize the data source %s",
dataSourceId));
        StartDataSourceSyncJobRequest startDataSourceSyncJobRequest =
StartDataSourceSyncJobRequest
            .builder()
            .indexId(indexId)
            .id(dataSourceId)
            .build();
        StartDataSourceSyncJobResponse startDataSourceSyncJobResponse =
kendra.startDataSourceSyncJob(startDataSourceSyncJobRequest);
        System.out.println(String.format("Waiting for the data
source to sync with the index %s for execution ID %s", indexId,
startDataSourceSyncJobResponse.executionId()));

        // For this example, there should be one job
```

```
        ListDataSourceSyncJobsRequest listDataSourceSyncJobsRequest =
 ListDataSourceSyncJobsRequest
            .builder()
            .indexId(indexId)
            .id(dataSourceId)
            .build();

        while (true) {
            ListDataSourceSyncJobsResponse listDataSourceSyncJobsResponse =
 kendra.listDataSourceSyncJobs(listDataSourceSyncJobsRequest);
            DataSourceSyncJob job = listDataSourceSyncJobsResponse.history().get(0);
            System.out.println(String.format("Syncing data source. Status: %s",
 job.status()));

            TimeUnit.SECONDS.sleep(60);
            if (job.status() != DataSourceSyncJobStatus.SYNCING) {
                break;
            }

        }

        System.out.println("Data source creation with customizations is complete");
    }
}
```

Contrats de données pour les fonctions Lambda

Vos fonctions Lambda pour la manipulation avancée des données interagissent avec les contrats de Amazon Kendra données. Les contrats sont les structures de demande et de réponse obligatoires de vos fonctions Lambda. Si vos fonctions Lambda ne suivent pas ces structures, une erreur est Amazon Kendra générée.

Votre fonction Lambda pour `PreExtractionHookConfiguration` devrait s'attendre à la structure de requête suivante :

```
{
    "version": <str>,
    "dataBlobStringEncodedInBase64": <str>, //In the case of a data blob
    "s3Bucket": <str>, //In the case of an S3 bucket
    "s3ObjectKey": <str>, //In the case of an S3 bucket
    "metadata": <Metadata>
```

```
}
```

La `metadata` structure, qui inclut la `CustomDocumentAttribute` structure, est la suivante :

```
{
    "attributes": [<CustomDocumentAttribute<]
}

CustomDocumentAttribute
{
    "name": <str>,
    "value": <CustomDocumentAttributeValue>
}

CustomDocumentAttributeValue
{
    "stringValue": <str>,
    "integerValue": <int>,
    "longValue": <long>,
    "stringListValue": list<str>,
    "dateValue": <str>
}
```

Votre fonction Lambda pour `PreExtractionHookConfiguration` doit respecter la structure de réponse suivante :

```
{
    "version": <str>,
    "dataBlobStringEncodedInBase64": <str>, //In the case of a data blob
    "s3ObjectKey": <str>, //In the case of an S3 bucket
    "metadataUpdates": [<CustomDocumentAttribute>]
}
```

Votre fonction Lambda pour `PostExtractionHookConfiguration` devrait s'attendre à la structure de requête suivante :

```
{
    "version": <str>,
    "s3Bucket": <str>,
    "s3ObjectKey": <str>,
    "metadata": <Metadata>
```

```
}
```

Votre fonction Lambda pour `PostExtractionHookConfiguration` doit respecter la structure de réponse suivante :

```
PostExtractionHookConfiguration Lambda Response
{
    "version": <str>,
    "s3ObjectKey": <str>,
    "metadataUpdates": [<CustomDocumentAttribute>]
}
```

Votre document modifié est chargé dans votre Amazon S3 compartiment. Le document modifié doit suivre le format indiqué dansthe section called “Format de document structuré”.

Format de document structuré

Amazon Kendra télécharge votre document structuré dans le Amazon S3 compartiment indiqué. Le document structuré suit le format suivant :

```
Kendra document

{
   "textContent": <TextContent>
}

TextContent
{
  "documentBodyText": <str>
}
```

Exemple de fonction Lambda qui respecte les contrats de données

Le code Python suivant est un exemple de fonction Lambda qui applique une manipulation avancée des champs `_authors` de métadonnées et du contenu du corps des documents bruts ou originaux. `_document_title`

Dans le cas où le contenu corporel se trouve dans un Amazon S3 compartiment

```
import json
```

```
import boto3

s3 = boto3.client("s3")

# Lambda function for advanced data manipulation
def lambda_handler(event, context):
    # Get the value of "S3Bucket" key name or item from the given event input
    s3_bucket = event.get("s3Bucket")
    # Get the value of "S3ObjectKey" key name or item from the given event input
    s3_object_key = event.get("s3ObjectKey")

    content_object_before_CDE = s3.get_object(Bucket = s3_bucket, Key = s3_object_key)
    content_before_CDE = content_object_before_CDE["Body"].read().decode("utf-8");
    content_after_CDE = "CDEInvolved " + content_before_CDE

    # Get the value of "metadata" key name or item from the given event input
    metadata = event.get("metadata")
    # Get the document "attributes" from the metadata
    document_attributes = metadata.get("attributes")

    s3.put_object(Bucket = s3_bucket, Key = "dummy_updated_kendra_document",
 Body=json.dumps(content_after_CDE))
    return {
        "version": "v0",
        "s3ObjectKey": "dummy_updated_kendra_document",
        "metadataUpdates": [
            {"name":"_document_title", "value":
{"stringValue":"title_from_pre_extraction_lambda"}},
            {"name":"_authors", "value":{"stringListValue":["author1", "author2"]}}
        ]
    }
```

Dans le cas où le contenu du corps réside dans un blob de données

```
import json
import boto3
import base64

# Lambda function for advanced data manipulation
def lambda_handler(event, context):

    # Get the value of "dataBlobStringEncodedInBase64" key name or item from the given
 event input
```

```
    data_blob_string_encoded_in_base64 = event.get("dataBlobStringEncodedInBase64")
    # Decode the data blob string in UTF-8
    data_blob_string =
 base64.b64decode(data_blob_string_encoded_in_base64).decode("utf-8")
    # Get the value of "metadata" key name or item from the given event input
    metadata = event.get("metadata")
    # Get the document "attributes" from the metadata
    document_attributes = metadata.get("attributes")

    new_data_blob = "This should be the modified data in the document by pre processing
 lambda ".encode("utf-8")
    return {
        "version": "v0",
        "dataBlobStringEncodedInBase64":
 base64.b64encode(new_data_blob).decode("utf-8"),
        "metadataUpdates": [
            {"name":"_document_title", "value":
{"stringValue":"title_from_pre_extraction_lambda"}},
            {"name":"_authors", "value":{"stringListValue":["author1", "author2"]}}
        ]
    }
```

Le code Python suivant est un exemple de fonction Lambda qui applique une manipulation avancée des champs `_authors` de métadonnées et du contenu du corps des documents structurés ou analysés. `_document_title`

```
import json
import boto3
import time

s3 = boto3.client("s3")

# Lambda function for advanced data manipulation
def lambda_handler(event, context):

    # Get the value of "S3Bucket" key name or item from the given event input
    s3_bucket = event.get("s3Bucket")
    # Get the value of "S3ObjectKey" key name or item from the given event input
    s3_key = event.get("s3ObjectKey")
    # Get the value of "metadata" key name or item from the given event input
    metadata = event.get("metadata")
    # Get the document "attributes" from the metadata
    document_attributes = metadata.get("attributes")
```

```
   kendra_document_object = s3.get_object(Bucket = s3_bucket, Key = s3_key)
   kendra_document_string = kendra_document_object['Body'].read().decode('utf-8')
   kendra_document = json.loads(kendra_document_string)
   kendra_document["textContent"]["documentBodyText"] = "Changing document body to a
short sentence."

   s3.put_object(Bucket = s3_bucket, Key = "dummy_updated_kendra_document",
Body=json.dumps(kendra_document))

   return {
       "version" : "v0",
       "s3ObjectKey": "dummy_updated_kendra_document",
       "metadataUpdates": [
           {"name": "_document_title", "value":{"stringValue":
"title_from_post_extraction_lambda"}},
           {"name": "_authors", "value":{"stringListValue":["author1", "author2"]}}
       ]
   }
```

Recherche dans un index

> **Note**
>
> La prise en charge des fonctionnalités varie en fonction du type d'index et de l'API de recherche utilisés. Pour savoir si cette fonctionnalité est prise en charge pour le type d'index et l'API de recherche que vous utilisez, consultez la section Types d'index.

Pour effectuer une recherche dans un Amazon Kendra index, vous utilisez l'API Query. L'`Query`API renvoie des informations sur les documents indexés que vous utilisez dans votre application. Cette section explique comment effectuer une requête, exécuter des filtres et interpréter la réponse que vous obtenez de l'`Query`API.

Pour rechercher des documents que vous avez indexés avec Amazon Kendra for Amazon Lex, utilisez AMAZON. KendraSearchIntent. Pour un exemple de configuration Amazon Kendra avec Amazon Lex, voir Création d'un bot FAQ pour un Amazon Kendra index.

Rubriques

- Interrogation d'un index
- Récupération de passages
- Parcourir un index
- Avec les résultats de recherche
- Recherche tabulaire pour le HTML
- suggestions de requêtes
- Correcteur orthographique des requêtes
- Filtrage et recherche par facettes
- Filtrage en fonction du contexte utilisateur
- Réponses aux requêtes et types de réponses
- Réglage et tri des réponses
- Réduction/extension des résultats de requête

Interrogation d'un index

> **Note**
>
> La prise en charge des fonctionnalités varie en fonction du type d'index et de l'API de recherche utilisés. Pour savoir si cette fonctionnalité est prise en charge pour le type d'index et l'API de recherche que vous utilisez, consultez la section Types d'index.

Lorsque vous effectuez une recherche dans votre index, il Amazon Kendra utilise toutes les informations que vous avez fournies sur vos documents pour déterminer les documents les plus pertinents par rapport aux termes de recherche saisis. Certains des éléments pris en Amazon Kendra compte sont les suivants :

- Le texte ou le corps du document.
- Titre du document.
- Champs de texte personnalisés que vous avez marqués comme consultables.
- Le champ de date que vous avez indiqué doit être utilisé pour déterminer la « fraîcheur » d'un document.
- Tout autre champ susceptible de fournir des informations pertinentes.

Amazon Kendra peut également filtrer la réponse en fonction des filtres de champ/d'attribut que vous pourriez avoir définis pour la recherche. Par exemple, si vous avez un champ personnalisé appelé « département », vous pouvez filtrer la réponse pour ne renvoyer que les documents provenant d'un département appelé « juridique ». Pour plus d'informations, consultez la section Champs ou attributs personnalisés.

Les résultats de recherche renvoyés sont triés en fonction de la pertinence qui Amazon Kendra détermine chaque document. Les résultats sont paginés afin que vous puissiez montrer une page à la fois à votre utilisateur.

Pour rechercher des documents que vous avez indexés avec Amazon Kendra for Amazon Lex, utilisez AMAZON. KendraSearchIntent. Pour un exemple de configuration Amazon Kendra avec Amazon Lex, voir Création d'un bot FAQ pour un Amazon Kendra index.

L'exemple suivant montre comment effectuer une recherche dans un index. Amazon Kendra détermine le type de résultat de recherche (réponse, document, question-réponse) le mieux adapté

à la requête. Vous ne pouvez pas configurer Amazon Kendra pour renvoyer un type spécifique de réponse de recherche (réponse, document, question-réponse) à une requête.

Pour plus d'informations sur les réponses aux requêtes, consultezRéponses aux requêtes et types de réponses.

Rubriques

- Prérequis
- Recherche dans un index (console)
- Recherche dans un index (SDK)
- Recherche dans un index (Postman)
- Recherche à l'aide d'une syntaxe de requête avancée
- Recherche dans les langues

Prérequis

Avant d'utiliser l'API Query pour interroger un index :

- Configurez les autorisations requises pour un index et connectez-vous à votre source de données ou téléchargez vos documents par lots. Pour plus d'informations, consultez la section IAM Rôles. Vous utilisez le nom de ressource Amazon du rôle lorsque vous appelez l'API pour créer un index et un connecteur de source de données ou pour télécharger des documents par lots.
- Configurez un SDK ou accédez à la Amazon Kendra console. AWS Command Line Interface Pour plus d’informations, consultez Configuration de Amazon Kendra.
- Créez un index et connectez-vous à une source de données de documents ou téléchargez des documents par lots. Pour plus d'informations, consultez les sections Création d'un index et Création d'un connecteur de source de données.

Recherche dans un index (console)

Vous pouvez utiliser la Amazon Kendra console pour rechercher et tester votre index. Vous pouvez effectuer des requêtes et voir les résultats.

Pour rechercher un index à l'aide de la console

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la à l'adresse http://console.aws.amazon.com/kendra/.
2. Dans le volet de navigation, sélectionnez Indexes.
3. Choisissez votre index.
4. Dans le menu de navigation, choisissez l'option permettant de rechercher dans votre index.
5. Entrez une requête dans la zone de texte, puis appuyez sur Entrée.
6. Amazon Kendra renvoie les résultats de la recherche.

Vous pouvez également obtenir l'ID de requête pour la recherche en sélectionnant l'icône représentant une ampoule dans le panneau latéral.

Recherche dans un index (SDK)

Pour rechercher un index avec Python ou Java

- L'exemple suivant effectue une recherche dans un index. Modifiez la valeur `query` de votre requête de recherche `index_id` et/ou `indexId` de l'identifiant d'index de l'index que vous souhaitez rechercher.

  Vous pouvez également obtenir l'ID de requête pour la recherche dans le cadre des éléments de réponse lorsque vous appelez l'API Query.

  Python

```
import boto3
import pprint

kendra = boto3.client("kendra")

# Provide the index ID
index_id = "index-id"
# Provide the query text
query = "query text"

response = kendra.query(
        QueryText = query,
        IndexId = index_id)
```

```
print("\nSearch results for query: " + query + "\n")

for query_result in response["ResultItems"]:

    print("-------------------")
    print("Type: " + str(query_result["Type"]))

    if query_result["Type"]=="ANSWER" or
 query_result["Type"]=="QUESTION_ANSWER":
        answer_text = query_result["DocumentExcerpt"]["Text"]
        print(answer_text)

    if query_result["Type"]=="DOCUMENT":
        if "DocumentTitle" in query_result:
            document_title = query_result["DocumentTitle"]["Text"]
            print("Title: " + document_title)
        document_text = query_result["DocumentExcerpt"]["Text"]
        print(document_text)

    print("------------------\n\n")
```

Java

```
package com.amazonaws.kendra;

import software.amazon.awssdk.services.kendra.KendraClient;
import software.amazon.awssdk.services.kendra.model.QueryRequest;
import software.amazon.awssdk.services.kendra.model.QueryResponse;
import software.amazon.awssdk.services.kendra.model.QueryResultItem;

public class SearchIndexExample {
    public static void main(String[] args) {
        KendraClient kendra = KendraClient.builder().build();

        String query = "query text";
        String indexId = "index-id";

        QueryRequest queryRequest = QueryRequest
            .builder()
            .queryText(query)
            .indexId(indexId)
```

```
            .build();

        QueryResponse queryResponse = kendra.query(queryRequest);

        System.out.println(String.format("\nSearch results for query: %s",
 query));
        for(QueryResultItem item: queryResponse.resultItems()) {
            System.out.println("----------------------");
            System.out.println(String.format("Type: %s", item.type()));

            switch(item.type()) {
                case QUESTION_ANSWER:
                case ANSWER:
                    String answerText = item.documentExcerpt().text();
                    System.out.println(answerText);
                    break;
                case DOCUMENT:
                    String documentTitle = item.documentTitle().text();
                    System.out.println(String.format("Title: %s",
 documentTitle));
                    String documentExcerpt = item.documentExcerpt().text();
                    System.out.println(String.format("Excerpt: %s",
 documentExcerpt));
                    break;
                default:
                    System.out.println(String.format("Unknown query result type:
 %s", item.type()));

            }

            System.out.println("----------------------\n");
        }
    }
}
```

Recherche dans un index (Postman)

Vous pouvez utiliser Postman pour interroger et tester votre Amazon Kendra index.

Pour effectuer une recherche dans un index à l'aide de Postman

1. Créez une nouvelle collection dans Postman et définissez le type de demande sur POST.

2. Entrez l'URL du point de terminaison. Par exemple, https://kendra. .amazonaws.com. <region>
3. Sélectionnez l'onglet Autorisation et entrez les informations suivantes.

   - Type —Sélectionnez une AWS signature.
   - AccessKey—Entrez la clé d'accès générée lors de la création d'un IAM utilisateur.
   - SecretKey—Entrez la clé secrète générée lors de la création d'un IAM utilisateur.
   - AWS Région —Entrez la région de votre index. Par exemple, us-west-2.
   - Nom du service —Entrez kendra. Ceci distingue les majuscules et minuscules, donc ce doit être fait en minuscules.

   > **Warning**
   >
   > Si vous entrez un nom de service incorrect ou si vous n'utilisez pas de minuscules, une erreur est générée lorsque vous sélectionnez Envoyer pour envoyer la demande : « Les informations d'identification doivent être limitées au bon service « kendra ». » Vous devez également vérifier que vous avez saisi la bonne clé d'accès et la bonne clé secrète.

4. Sélectionnez l'onglet En-têtes et entrez les informations de clé et de valeur suivantes.

   - Clé : X-Amz-Target

     Valeur : com.amazonaws.kendra. AWSKendraFrontendService.Requête
   - Clé : Encodage du contenu

     Valeur : amz-1.0
5. Sélectionnez l'onglet Body et procédez comme suit.

   - Choisissez le type JSON brut pour le corps de la demande.
   - Entrez un JSON qui inclut votre identifiant d'index et le texte de votre requête.

   ```
   {
       "IndexId": "index-id",
       "QueryText": "enter a query here"
   }
   ```

> **Warning**
>
> Si votre JSON n'utilise pas l'indendation correcte, une erreur est renvoyée : « »SerializationException. Vérifiez l'indendation dans votre JSON.

6. Sélectionnez Envoyer (en haut à droite).

Recherche à l'aide d'une syntaxe de requête avancée

Vous pouvez créer des requêtes plus spécifiques que de simples requêtes par mots clés ou en langage naturel en utilisant une syntaxe ou des opérateurs de requête avancés. Cela inclut les plages, les booléens, les caractères génériques, etc. En utilisant des opérateurs, vous pouvez donner plus de contexte à votre requête et affiner les résultats de recherche.

Amazon Kendra prend en charge les opérateurs suivants.

- Booléen : logique permettant de limiter ou d'élargir la recherche. Par exemple, `amazon AND sports` limite la recherche aux seuls documents contenant les deux termes.
- Parenthèses : lit les termes de requête imbriqués par ordre de priorité. Par exemple, `(amazon AND sports) NOT rainforest` lit `(amazon AND sports)` avant`NOT rainforest`.
- Plages : dates ou valeurs de plage numériques. Les plages peuvent être inclusives, exclusives ou illimitées. Par exemple, vous pouvez rechercher des documents qui ont été mis à jour pour la dernière fois entre le 1er janvier 2020 et le 31 décembre 2020, y compris ces dates.
- Champs : Utilise un champ spécifique pour limiter la recherche. Par exemple, vous pouvez rechercher des documents dont le champ « localisation » contient « États-Unis ».
- Caractères génériques : correspondent partiellement à une chaîne de texte. Par exemple, `Cloud*` pourrait correspondre CloudFormation. Amazon Kendra ne prend actuellement en charge que les caractères génériques de fin.
- Guillemets exacts : correspondent exactement à une chaîne de texte. Par exemple, les documents contenant`"Amazon Kendra" "pricing"`.

Vous pouvez utiliser une combinaison de n'importe lequel des opérateurs ci-dessus.

Notez qu'une utilisation excessive d'opérateurs ou des requêtes très complexes peuvent avoir un impact sur la latence des requêtes. Les Wildcards font partie des opérateurs les plus chers en termes de latence. En règle générale, plus vous utilisez de termes et d'opérateurs, plus l'impact potentiel sur

la latence est important. Parmi les autres facteurs qui influent sur la latence, citons la taille moyenne des documents indexés, la taille de votre index, les éventuels filtres appliqués aux résultats de recherche et la charge globale de votre Amazon Kendra index.

Booléen

Vous pouvez combiner ou exclure des mots à l'aide des opérateurs booléens`AND`,,`OR`. `NOT`

Voici des exemples d'utilisation d'opérateurs booléens.

`amazon AND sports`

Renvoie les résultats de recherche qui contiennent à la fois les termes « amazon » et « sports » dans le texte, tels que Amazon Prime Video Sports ou tout autre contenu similaire.

`sports OR recreation`

Renvoie les résultats de recherche qui contiennent les termes « sports » ou « loisirs », ou les deux, dans le texte.

`amazon NOT rainforest`

Renvoie les résultats de recherche qui contiennent le terme « amazone » mais pas le terme « forêt tropicale » dans le texte. Il s'agit de rechercher des documents sur l'entreprise Amazon, et non sur la forêt amazonienne.

Parenthèses

Vous pouvez interroger les mots imbriqués par ordre de priorité en utilisant des parenthèses. Les parenthèses indiquent Amazon Kendra comment une requête doit être lue.

Voici des exemples d'utilisation d'opérateurs entre parenthèses.

`(amazon AND sports) NOT rainforest`

Renvoie les documents qui contiennent à la fois les termes « amazon » et « sports » dans le texte, mais pas le terme « forêt tropicale ». Il s'agit de rechercher des vidéos sur Amazon Prime, des sports ou d'autres contenus similaires, et non des sports d'aventure dans la forêt amazonienne. Les parenthèses aident à indiquer que cela `amazon AND sports` doit être lu avant`NOT rainforest`. La requête ne doit pas être lue comme telle`amazon AND (sports NOT rainforest)`.

`(amazon AND (sports OR recreation)) NOT rainforest`

Renvoie les documents contenant les termes « sport » ou « loisirs », ou les deux, et le terme « Amazon ». Mais le terme « forêt tropicale » n'y figure pas. Il s'agit de rechercher des vidéos sportives ou récréatives sur Amazon Prime, et non des sports d'aventure dans la forêt amazonienne. Les parenthèses indiquent qu'il `sports OR recreation` faut le lire avant de le combiner avec « amazon », qui est lu avant`NOT rainforest`. La requête ne doit pas être lue comme telle`amazon AND (sports OR (recreation NOT rainforest))`.

Gammes

Vous pouvez utiliser une plage de valeurs pour filtrer les résultats de recherche. Vous spécifiez un attribut et les valeurs de plage. Il peut s'agir d'une date ou d'un type numérique.

Les plages de dates doivent respecter les formats suivants :

- Epoch
- YYYY
- YYYY-mm
- YYYY-MM-DD
- YYYY-MM-DD'T'HH

Vous pouvez également spécifier s'il faut inclure ou exclure les valeurs inférieures et supérieures de la plage.

Vous trouverez ci-dessous des exemples d'utilisation des opérateurs de plage.

`_processed_date:>2019-12-31 AND _processed_date:<2021-01-01`

Renvoie les documents qui ont été traités en 2020, plus tard que le 31 décembre 2019 et moins que le 1er janvier 2021.

`_processed_date:>=2020-01-01 AND _processed_date:<=2020-12-31`

Renvoie les documents traités en 2020, supérieurs ou égaux au 1er janvier 2020 et inférieurs ou égaux au 31 décembre 2020.

`_document_likes:<1`

Renvoie les documents n'ayant reçu aucun « j'aime » ou aucun commentaire de la part de l'utilisateur, c'est-à-dire moins d'un « j'aime ».

Vous pouvez spécifier si une plage doit être traitée comme incluant ou excluant les valeurs de plage données.

Inclusif

`_last_updated_at:[2020-01-01 TO 2020-12-31]`

Les documents de retour ont été mis à jour pour la dernière fois en 2020, y compris les jours du 1er décembre 2020 et du 31 décembre 2020.

Exclusif

`_last_updated_at:{2019-12-31 TO 2021-01-01}`

Les documents de retour ont été mis à jour pour la dernière fois en 2020, à l'exception des jours du 31 décembre 2019 et du 1er janvier 2021.

< and **>**Pour les plages illimitées qui ne sont ni inclusives ni exclusives, utilisez simplement les opérateurs. Par exemple, `_last_updated_at:>2019-12-31 AND _last_updated_at:<2021-01-01`

Champs

Vous pouvez limiter votre recherche pour ne renvoyer que les documents correspondant à une valeur dans un champ spécifique. Le champ peut être de n'importe quel type.

Vous trouverez ci-dessous des exemples d'utilisation d'opérateurs de contexte au niveau du champ.

`status:"Incomplete" AND financial_year:2021`

Renvoie les documents de l'exercice 2021 avec leur statut incomplet.

`(sports OR recreation) AND country:"United States" AND level:"professional"`

Renvoie les documents qui traitent des sports professionnels ou des loisirs aux États-Unis d'Amérique.

Caractères génériques

Vous pouvez élargir votre recherche pour prendre en compte les variantes de mots et d'expressions à l'aide de l'opérateur générique. Cela est utile lors de la recherche de variantes de noms. Amazon Kendra ne prend actuellement en charge que les caractères génériques de fin. Le nombre de caractères de préfixe pour un caractère générique final doit être supérieur à deux.

Voici des exemples d'utilisation d'opérateurs génériques.

`Cloud*`

Renvoie les documents contenant des variantes telles que CloudFormation et CloudWatch.

`kendra*aws`

Renvoie les documents contenant des variantes telles que kendra.amazonaws.

`kendra*aws*`

Renvoie les documents contenant des variantes, tels que kendra.amazonaws.com

Citations exactes

Vous pouvez utiliser des guillemets pour rechercher une correspondance exacte entre un texte.

Vous trouverez ci-dessous des exemples d'utilisation de guillemets.

`"Amazon Kendra" "pricing"`

Renvoie les documents contenant à la fois l'expression « Amazon Kendra » et le terme « tarification ». Les documents doivent contenir à la fois des Amazon Kendra« » et des « prix » pour que les résultats puissent être renvoyés.

`"Amazon Kendra" "pricing" cost`

Renvoie les documents contenant à la fois l'expression « Amazon Kendra » et le terme « tarification », et éventuellement le terme « coût ». Les documents doivent contenir à la fois le terme « Amazon Kendra » et le « prix » afin de fournir les résultats, mais ils ne doivent pas nécessairement inclure le « coût ».

Syntaxe de requête non valide

Amazon Kendra émet un avertissement en cas de problème de syntaxe de votre requête ou si votre requête n'est actuellement pas prise en charge par Amazon Kendra. Pour plus d'informations, consultez la documentation de l'API pour les avertissements relatifs aux requêtes.

Les requêtes suivantes sont des exemples de syntaxe de requête non valide.

`_last_updated_at:<2021-12-32`

Date non valide. Le jour 32 n'existe pas dans le calendrier grégorien, qui est utilisé par Amazon Kendra.

`_view_count:ten`

Valeur numérique non valide. Les chiffres doivent être utilisés pour représenter des valeurs numériques.

`nonExistentField:123`

Recherche par champ non valide. Le champ doit exister pour pouvoir utiliser la recherche par champ.

`Product:[A TO D]`

Plage non valide. Des valeurs numériques ou des dates doivent être utilisées pour les plages.

`OR Hello`

Booléen non valide. Les opérateurs doivent être utilisés avec des termes et placés entre les termes.

Recherche dans les langues

Vous pouvez rechercher des documents dans une langue prise en charge. Vous devez saisir le code de langue AttributeFilterpour renvoyer les documents filtrés dans la langue de votre choix. Vous pouvez saisir la requête dans une langue prise en charge.

Si vous ne spécifiez aucune langue, Amazon Kendra interroge les documents en anglais par défaut. Pour plus d'informations sur les langues prises en charge, y compris leurs codes, voir Ajout de documents dans des langues autres que l'anglais.

Pour rechercher des documents dans une langue prise en charge dans la console, sélectionnez votre index, puis sélectionnez l'option permettant de rechercher dans votre index dans le menu de navigation. Choisissez la langue dans laquelle vous souhaitez renvoyer les documents en sélectionnant les paramètres de recherche, puis en sélectionnant une langue dans le menu déroulant Langue.

Les exemples suivants montrent comment rechercher des documents en espagnol.

Pour rechercher un index en espagnol dans la console

1. Connectez-vous à la Amazon Kendra console AWS Management Console et ouvrez-la à l'adresse http://console.aws.amazon.com/kendra/.
2. Dans le menu de navigation, choisissez Index et choisissez votre index.

3. Dans le menu de navigation, choisissez l'option permettant de rechercher dans votre index.
4. Dans les paramètres de recherche, sélectionnez le menu déroulant Langues et choisissez Espagnol.
5. Entrez une requête dans la zone de texte, puis appuyez sur Entrée.
6. Amazon Kendra renvoie les résultats de la recherche en espagnol.

Pour rechercher un index en espagnol à l'aide de la CLI, de Python ou de Java

- L'exemple suivant effectue une recherche dans un index en espagnol. Modifiez la valeur `searchString` de votre requête de recherche et la valeur `indexID` de l'identifiant de l'index que vous souhaitez rechercher. Le code de langue pour l'espagnol est`es`. Vous pouvez le remplacer par votre propre code de langue.

  CLI

  ```
  {
    "EqualsTo":{
      "Key": "_language_code",
      "Value": {
      "StringValue": "es"
      }
    }
  }
  ```

  Python

  ```
  import boto3
  import pprint

  kendra = boto3.client("kendra")

  # Provide the index ID
  index_id = "index-id"
  # Provide the query text
  query = "search-string"

  # Includes the index ID, query text, and language attribute filter
  response = kendra.query(
          QueryText = query,
          IndexId = index_id,
  ```

```
        AttributeFilter = {
            "EqualsTo": {
                "Key": "_language_code",
                "Value": {
                    "StringValue": "es"
                    }
                }
            })

print ("\nSearch results|Resultados de la búsqueda: " + query + "\n")

for query_result in response["ResultItems"]:

    print("-------------------")
    print("Type: " + str(query_result["Type"]))

    if query_result["Type"]=="ANSWER" or
 query_result["Type"]=="QUESTION_ANSWER":
        answer_text = query_result["DocumentExcerpt"]["Text"]
        print(answer_text)

    if query_result["Type"]=="DOCUMENT":
        if "DocumentTitle" in query_result:
            document_title = query_result["DocumentTitle"]["Text"]
            print("Title: " + document_title)
        document_text = query_result["DocumentExcerpt"]["Text"]
        print(document_text)

    print("------------------\n\n")
```

Java

```
package com.amazonaws.kendra;

import software.amazon.awssdk.services.kendra.KendraClient;
import software.amazon.awssdk.services.kendra.model.QueryRequest;
import software.amazon.awssdk.services.kendra.model.QueryResponse;
import software.amazon.awssdk.services.kendra.model.QueryResultItem;

public class SearchIndexExample {
    public static void main(String[] args) {
```

```
        KendraClient kendra = KendraClient.builder().build();

        String query = "searchString";
        String indexId = "indexID";

        QueryRequest queryRequest = QueryRequest.builder()
            .queryText(query)
            .indexId(indexId)
            .attributeFilter(
                 AttributeFilter.builder()
                     .withEqualsTo(
                         DocumentAttribute.builder()
                             .withKey("_language_code")
                             .withValue("es")
                             .build())
                     .build())
            .build();

        QueryResponse queryResponse = kendra.query(queryRequest);

        System.out.println(String.format("\nSearch results|
                                          Resultados de la búsqueda: %s",
query));
        for(QueryResultItem item: queryResponse.resultItems()) {
            System.out.println("----------------------");
            System.out.println(String.format("Type: %s", item.type()));

            switch(item.type()) {
                case QUESTION_ANSWER:
                case ANSWER:
                    String answerText = item.documentExcerpt().text();
                    System.out.println(answerText);
                    break;
                case DOCUMENT:
                    String documentTitle = item.documentTitle().text();
                    System.out.println(String.format("Title: %s",
documentTitle));
                    String documentExcerpt = item.documentExcerpt().text();
                    System.out.println(String.format("Excerpt: %s",
documentExcerpt));
                    break;
                default:
                    System.out.println(String.format("Unknown query result type:
%s", item.type()));
```

```
            }

            System.out.println("----------------------\n");
        }
    }
}
```

Récupération de passages

Vous pouvez utiliser le RetrieveAPI en tant que récupérateur pour les systèmes de génération augmentée (RAG).

Les systèmes RAG utilisent l'intelligence artificielle générative pour créer des applications de réponse aux questions. Les systèmes RAG se composent d'un récupérateur et de grands modèles linguistiques (LLM). À la suite d'une requête, le récupérateur identifie les parties de texte les plus pertinentes d'un corpus de documents et les transmet au LLM pour fournir la réponse la plus utile. Ensuite, le LLM analyse les fragments ou passages de texte pertinents et génère une réponse complète à la requête.

L'`Retrieve`API examine les fragments de texte ou les extraits appelés passages et renvoie les premiers passages les plus pertinents pour la requête.

Comme le QueryAPI, l'`Retrieve`API recherche également les informations pertinentes. La récupération d'informations de l'API Retrieve prend en compte le contexte de la requête et toutes les informations disponibles à partir des documents indexés. Toutefois, par défaut, l'`Query`API ne renvoie que des extraits contenant jusqu'à 100 mots symboliques. Grâce à l'`Retrieve`API, vous pouvez récupérer des passages plus longs contenant jusqu'à 200 mots symboliques et jusqu'à 100 passages pertinents du point de vue sémantique. Cela n'inclut pas les réponses de type question-réponse ou FAQ de votre index. Les passages, également appelés fragments, sont des extraits de texte qui peuvent être extraits sémantiquement de plusieurs documents et de plusieurs parties d'un même document. L'indice GenAI Enterprise Edition de Kendra fournit des résultats de haute précision à récupérer, en utilisant une recherche hybride sur des indices de vecteurs et de mots clés, ainsi qu'un classement par modèles d'apprentissage en profondeur.

Vous pouvez également effectuer les opérations suivantes avec l'`Retrieve`API :

- Annuler le boost au niveau de l'indice
- Filtrer en fonction des champs ou des attributs du document

- Filtrer en fonction de l'accès de l'utilisateur ou de son groupe aux documents
- Consultez le compartiment des scores de confiance pour obtenir un résultat de passage récupéré. Le bucket de confiance fournit un classement relatif qui indique dans quelle mesure la réponse est pertinente par rapport à la requête. Amazon Kendra

> **Note**
>
> Les buckets de scores de confiance ne sont actuellement disponibles qu'en anglais.

Vous pouvez également inclure dans la réponse certains champs susceptibles de fournir des informations supplémentaires utiles.

L'`Retrieve`API ne prend actuellement pas en charge les fonctionnalités suivantes : requêtes utilisant une syntaxe de requête avancée, corrections orthographiques suggérées pour les requêtes, facettage, suggestions de requêtes pour compléter automatiquement les requêtes de recherche et apprentissage progressif. Aucune requête d'API de récupération n'apparaîtra dans le tableau de bord d'analyse.

L'`Retrieve`API partage le nombre d'unités de capacité de requête que vous définissez pour votre index. Pour plus d'informations sur ce qui est inclus dans une unité de capacité unique et sur la capacité de base par défaut d'un indice, consultez la section Ajustement de la capacité.

> **Note**
>
> Vous ne pouvez pas ajouter de capacité si vous utilisez l'édition Amazon Kendra Developer ; vous ne pouvez ajouter de la capacité que si vous utilisez l'édition Amazon Kendra Enterprise. Pour plus d'informations sur ce qui est inclus dans les éditions Developer et Enterprise, consultez la section Amazon Kendra Éditions.

Voici un exemple d'utilisation de l'`Retrieve`API pour récupérer les 100 passages les plus pertinents des documents dans un index pour la requête. "how does amazon kendra work?"

Python

```
import boto3
import pprint
```

```
kendra = boto3.client("kendra")

# Provide the index ID
index_id = "index-id"
# Provide the query text
query = "how does amazon kendra work?"
# You can retrieve up to 100 relevant passages
# You can paginate 100 passages across 10 pages, for example
page_size = 10
page_number = 10

result = kendra.retrieve(
        IndexId = index_id,
        QueryText = query,
        PageSize = page_size,
        PageNumber = page_number)

print("\nRetrieved passage results for query: " + query + "\n")

for retrieve_result in result["ResultItems"]:

    print("--------------------")
    print("Title: " + str(retrieve_result["DocumentTitle"]))
    print("URI: " + str(retrieve_result["DocumentURI"]))
    print("Passage content: " + str(retrieve_result["Content"]))
    print("------------------\n\n")
```

Java

```
package com.amazonaws.kendra;

import software.amazon.awssdk.services.kendra.KendraClient;
import software.amazon.awssdk.services.kendra.model.RetrieveRequest;
import software.amazon.awssdk.services.kendra.model.RetrieveResult;
import software.amazon.awssdk.services.kendra.model.RetrieveResultItem;

public class RetrievePassageExample {
    public static void main(String[] args) {
        KendraClient kendra = KendraClient.builder().build();

        String indxId = "index-id";
        String query = "how does amazon kendra work?";
        Integer pgSize = 10;
```

```
        Integer pgNumber = 10;

        RetrieveRequest retrieveRequest = retrieveRequest
            .builder()
            .indexId(indxId)
            .queryText(query)
            .pageSize(pgSize)
            .pageNumber(pgNumber)
            .build();

        RetrieveResult retrieveResult = kendra.retrieve(retrieveRequest);

        System.out.println(String.format("\nRetrieved passage results for query:
 %s", query));
        for(RetrieveResultItem item: retrieveResult.resultItems()) {
            System.out.println("----------------------");
            System.out.println(String.format("Title: %s", documentTitle));
            System.out.println(String.format("URI: %s", documentURI));
            System.out.println(String.format("Passage content: %s", content));
            System.out.println("----------------------\n");
        }
    }
}
```

Parcourir un index

> **Note**
>
> La prise en charge des fonctionnalités varie en fonction du type d'index et de l'API de recherche utilisés. Pour savoir si cette fonctionnalité est prise en charge pour le type d'index et l'API de recherche que vous utilisez, consultez la section Types d'index.

Vous pouvez parcourir les documents en fonction de leurs attributs ou de leurs facettes sans avoir à saisir de requête de recherche. Amazon Kendra Index Browse peut aider vos utilisateurs à découvrir des documents en parcourant librement un index sans avoir à répondre à une requête spécifique. Cela permet également à vos utilisateurs de parcourir largement un index comme point de départ de leur recherche.

Index Browse ne peut être utilisé que pour effectuer une recherche par attribut ou facette du document avec un type de tri. Vous ne pouvez pas effectuer de recherche dans un index complet à l'aide de l'outil Index Browse. Si le texte de la requête est absent, Amazon Kendra demande un filtre d'attribut de document ou une facette, ainsi qu'un type de tri.

Pour autoriser la navigation dans les index à l'aide de l'API Query, vous devez inclure AttributeFilterou Facet, et SortingConfiguration. Pour autoriser la navigation par index dans la console, sélectionnez votre index sous Index dans le menu de navigation, puis sélectionnez l'option permettant de rechercher dans votre index. Dans le champ de recherche, appuyez deux fois sur la touche Entrée. Sélectionnez le menu déroulant Filtrer les résultats de recherche pour choisir un filtre et sélectionnez le menu déroulant Trier pour choisir un type de tri.

Voici un exemple de navigation dans un index de documents en espagnol par ordre décroissant de date de création du document.

CLI

```
aws kendra query \
--index-id "index-id" \
--attribute-filter '{
    "EqualsTo":{
      "Key": "_language_code",
      "Value": {
        "StringValue": "es"
      }
    }
  }' \
--sorting-configuration '{
    "DocumentAttributeKey": "_created_at",
    "SortOrder": "DESC"
  }'
```

Python

```
import boto3

kendra = boto3.client("kendra")

# Must include the index ID, the attribute filter, and sorting configuration
response = kendra.query(
        IndexId = "index-id",
        AttributeFilter = {
```

```
            "EqualsTo": {
                "Key": "_language_code",
                "Value": {
                    "StringValue": "es"
                    }
                }
            },
        SortingConfiguration = {
            "DocumentAttributeKey": "_created_at",
            "SortOrder": "DESC"})

print("\nSearch results|Resultados de la búsqueda: \n")

for query_result in response["ResultItems"]:

    print("-------------------")
    print("Type: " + str(query_result["Type"]))

    if query_result["Type"]=="ANSWER" or query_result["Type"]=="QUESTION_ANSWER":
        answer_text = query_result["DocumentExcerpt"]["Text"]
        print(answer_text)

    if query_result["Type"]=="DOCUMENT":
        if "DocumentTitle" in query_result:
            document_title = query_result["DocumentTitle"]["Text"]
            print("Title: " + document_title)
        document_text = query_result["DocumentExcerpt"]["Text"]
        print(document_text)

    print("------------------\n\n")
```

Java

```
package com.amazonaws.kendra;

import software.amazon.awssdk.services.kendra.KendraClient;
import software.amazon.awssdk.services.kendra.model.QueryRequest;
import software.amazon.awssdk.services.kendra.model.QueryResult;
import software.amazon.awssdk.services.kendra.model.QueryResultItem;

public class SearchIndexExample {
    public static void main(String[] args) {
        KendraClient kendra = KendraClient.builder().build();
```

```
        QueryRequest queryRequest = QueryRequest.builder()
            .withIndexId("index-id")
            .withAttributeFilter(AttributeFilter.builder()
                .withEqualsTo(DocumentAttribute.builder()
                    .withKey("_language_code")
                    .withValue(DocumentAttributeValue.builder()
                        .withStringValue("es")
                        .build())
                    .build())
                .build())
            .withSortingConfiguration(SortingConfiguration.builder()
                .withDocumentAttributeKey("_created_at")
                .withSortOrder("DESC")
                .build())
            .build());

        QueryResult queryResult = kendra.query(queryRequest);
        for (QueryResultItem item : queryResult.getResultItems()) {
            System.out.println("----------------------");
            System.out.println(String.format("Type: %s", item.getType()));

            switch (item.getType()) {
                case QueryResultType.QUESTION_ANSWER:
                case QueryResultType.ANSWER:
                    String answerText = item.getDocumentExcerpt().getText();
                    System.out.println(answerText);
                    break;
                case QueryResultType.DOCUMENT:
                    String documentTitle = item.getDocumentTitle().getText();
                    System.out.println(String.format("Title: %s", documentTitle));
                    String documentExcerpt = item.getDocumentExcerpt().getText();
                    System.out.println(String.format("Excerpt: %s",
 documentExcerpt));
                    break;
                default:
                    System.out.println(String.format("Unknown query result type:
 %s", item.getType()));
            }
            System.out.println("----------------------\n");
        }
    }
}
```

Avec les résultats de recherche

> **Note**
>
> La prise en charge des fonctionnalités varie en fonction du type d'index et de l'API de recherche utilisés. Pour savoir si cette fonctionnalité est prise en charge pour le type d'index et l'API de recherche que vous utilisez, consultez la section Types d'index.

Vous pouvez faire apparaître certains documents dans les résultats de recherche lorsque vos utilisateurs émettent certaines requêtes. Cela permet de rendre les résultats plus visibles et plus visibles pour vos utilisateurs. Les résultats présentés sont séparés de la liste de résultats habituelle et affichés en haut de la page de recherche. Vous pouvez essayer de proposer différents documents pour différentes requêtes ou vous assurer que certains documents obtiennent la visibilité qu'ils méritent.

Vous associez des requêtes spécifiques à des documents spécifiques pour les inclure dans les résultats. Si une requête contient une correspondance exacte, un ou plusieurs documents spécifiques apparaissent dans les résultats de recherche.

Par exemple, vous pouvez spécifier que si vos utilisateurs émettent la requête « nouveaux produits 2023 », ils sélectionnent les documents intitulés « Nouveautés » et « Prochainement » à afficher en haut de la page des résultats de recherche. Cela permet de garantir que ces documents sur les nouveaux produits obtiennent la visibilité qu'ils méritent.

Amazon Kendra ne duplique pas les résultats de recherche si un résultat est déjà sélectionné pour figurer en haut de la page des résultats de recherche. Un résultat en vedette n'est pas à nouveau classé comme premier résultat s'il figure déjà au-dessus de tous les autres résultats.

Pour obtenir certains résultats, vous devez spécifier une correspondance exacte à une requête en texte intégral, et non une correspondance partielle à une requête utilisant un mot clé ou une phrase contenue dans une requête. Par exemple, si vous spécifiez uniquement la requête « Kendra » dans un ensemble de résultats en vedette, des requêtes telles que « Comment Kendra classe-t-elle sémantiquement les résultats ? » n'affichera pas les résultats présentés. Les résultats présentés sont conçus pour des requêtes spécifiques, plutôt que pour des requêtes dont la portée est trop large. Amazon Kendra gère naturellement les requêtes de type mot clé afin de classer les documents les plus utiles dans les résultats de recherche, évitant ainsi une présentation excessive des résultats basés sur des mots clés simples.

Si certaines requêtes sont fréquemment utilisées par vos utilisateurs, vous pouvez les spécifier pour obtenir des résultats en vedette. Par exemple, si vous examinez vos principales requêtes à l'aide Amazon Kendra d'Analytics et que vous trouvez des requêtes spécifiques, telles que « Comment Kendra classe-t-elle sémantiquement les résultats ? » et « Kendra Semantic Search » sont fréquemment utilisées. Il peut donc être utile de spécifier ces requêtes pour présenter le document intitulé « search 101 ».Amazon Kendra

Amazon Kendra traite les requêtes portant sur les résultats présentés sans distinction majuscules/minuscules. Amazon Kendra convertit une requête en minuscules et remplace les espaces blancs de fin par un seul espace. Amazon Kendra correspond à tous les autres caractères tels qu'ils sont lorsque vous spécifiez vos requêtes pour les résultats présentés.

Vous créez un ensemble de résultats en vedette que vous associez à certaines requêtes à l'aide de l'CreateFeaturedResultsSetAPI. Si vous utilisez la console, vous sélectionnez votre index, puis sélectionnez Résultats en vedette dans le menu de navigation pour créer un ensemble de résultats sélectionnés. Vous pouvez créer jusqu'à 50 ensembles de résultats en vedette par index, jusqu'à quatre documents à présenter par ensemble et jusqu'à 49 textes de requête par ensemble de résultats en vedette. Vous pouvez demander à augmenter ces limites en contactant le Support.

Vous pouvez sélectionner le même document parmi plusieurs ensembles de résultats présentés. Toutefois, vous ne devez pas utiliser le même texte de requête de correspondance exacte dans plusieurs ensembles. Les requêtes que vous spécifiez pour les résultats en vedette doivent être uniques par résultat en vedette défini pour chaque index.

Vous pouvez organiser l'ordre des documents lorsque vous sélectionnez jusqu'à quatre documents en vedette. Si vous utilisez l'API, l'ordre dans lequel vous listez les documents en vedette est le même que celui affiché dans les résultats présentés. Si vous utilisez la console, vous pouvez simplement glisser-déposer l'ordre des documents lorsque vous sélectionnez les documents à inclure dans les résultats.

Le contrôle d'accès, selon lequel certains utilisateurs et groupes ont accès à certains documents et d'autres non, est toujours respecté lors de la configuration des résultats présentés. Cela vaut également pour le filtrage du contexte utilisateur. Par exemple, l'utilisateur A appartient au groupe d'entreprises « Stagiaires », qui ne doit pas accéder aux documents relatifs aux secrets de l'entreprise. Si l'utilisateur A saisit une requête contenant un document secret d'entreprise, l'utilisateur A ne voit pas ce document apparaître dans ses résultats. Cela vaut également pour tous les autres résultats de la page des résultats de recherche. Vous pouvez également utiliser des balises pour contrôler l'accès à un ensemble de résultats en vedette, qui est une Amazon Kendra ressource pour laquelle vous contrôlez l'accès.

Voici un exemple de création d'un ensemble de résultats en vedette avec les requêtes « nouveaux produits 2023 », « nouveaux produits disponibles » mappées aux documents intitulés « Nouveautés » (doc-id-1) et « Prochainement » (doc-id-2).

CLI

```
aws kendra create-featured-results-set \
 --featured-results-set-name 'New product docs to feature' \
 --description "Featuring What's new and Coming soon docs" \
 --index-id index-id \
 --query-texts 'new products 2023' 'new products available' \
 --featured-documents '{"Id":"doc-id-1", "Id":"doc-id-2"}'
```

Python

```
import boto3
from botocore.exceptions import ClientError
import pprint
import time

kendra = boto3.client("kendra")

print("Create a featured results set.")

# Provide a name for the featured results set
featured_results_name = "New product docs to feature"
# Provide an optional decription for the featured results set
description = "Featuring What's new and Coming soon docs"
# Provide the index ID for the featured results set
index = "index-id"
# Provide a list of query texts for the featured results set
queries = ['new products 2023', 'new products available']
# Provide a list of document IDs for the featured results set
featured_doc_ids = [{"Id":"doc-id-1"}, {"Id":"doc-id-2"}]

try:
    featured_results_set_response = kendra.create_featured_results_set(
        FeaturedResultsSetName = featured_results_name,
        Decription = description,
        Index = index,
        QueryTexts = queries,
        FeaturedDocuments = featured_doc_ids
    )
```

```
    pprint.pprint(featured_results_set_response)

    featured_results_set_id = featured_results_set_response["FeaturedResultsSetId"]

    while True:
        # Get the details of the featured results set, such as the status
        featured_results_set_description = kendra.describe_featured_results_set(
            Id = featured_results_set_id
        )
        status = featured_results_set_description["Status"]
        print(" Featured results set status: "+status)

except  ClientError as e:
        print("%s" % e)

print("Program ends.")
```

Recherche tabulaire pour le HTML

> **Note**
>
> La prise en charge des fonctionnalités varie en fonction du type d'index et de l'API de recherche utilisés. Pour savoir si cette fonctionnalité est prise en charge pour le type d'index et l'API de recherche que vous utilisez, consultez la section Types d'index.

Amazon Kendra La fonction de recherche tabulaire permet de rechercher et d'extraire des réponses à partir de tableaux intégrés dans des documents HTML. Lorsque vous effectuez une recherche dans votre index, Amazon Kendra inclut un extrait d'un tableau s'il est pertinent pour la requête et fournit des informations utiles.

Amazon Kendra examine toutes les informations contenues dans le corps du texte d'un document, y compris les informations utiles contenues dans les tableaux. Par exemple, un index contient des rapports commerciaux avec des tableaux sur les coûts d'exploitation, les revenus et d'autres informations financières. Pour la question, « quel est le coût de fonctionnement annuel de 2020 à 2022 ? » , Amazon Kendra peut renvoyer un extrait d'un tableau contenant les colonnes pertinentes « Opérations (millions de dollars américains) » et « Exercice financier », ainsi que des lignes de tableau contenant les valeurs des revenus pour 2020, 2021 et 2022. L'extrait du tableau est inclus

dans le résultat, ainsi que le titre du document, un lien vers le document complet et tout autre champ du document que vous choisissez d'inclure.

Des extraits de tableau peuvent être affichés dans les résultats de recherche, que les informations se trouvent dans une ou plusieurs cellules d'un tableau. Par exemple, Amazon Kendra vous pouvez afficher un extrait de tableau adapté à chacun de ces types de requêtes :

- « carte de crédit au taux d'intérêt le plus élevé en 2020 »
- « carte de crédit au taux d'intérêt le plus élevé de 2020-2022"
- « Les 3 cartes de crédit ayant le taux d'intérêt le plus élevé en 2020-2022"
- « cartes de crédit dont le taux d'intérêt est inférieur à 10 % »
- « toutes les cartes de crédit à faible taux d'intérêt disponibles »

Amazon Kendra met en évidence la ou les cellules du tableau les plus pertinentes pour la requête. Les cellules les plus pertinentes avec leurs lignes, colonnes et noms de colonnes correspondants sont affichées dans les résultats de recherche. L'extrait du tableau affiche jusqu'à cinq colonnes et trois lignes, selon le nombre de cellules du tableau correspondant à la requête et le nombre de colonnes disponibles dans le tableau d'origine. La cellule la plus pertinente du haut est affichée dans l'extrait du tableau, ainsi que les cellules les plus pertinentes suivantes.

La réponse inclut le compartiment de confiance (`MEDIUM`,`HIGH`,`VERY_HIGH`) pour montrer dans quelle mesure la réponse de la table est pertinente par rapport à la requête. Si la valeur d'une cellule de tableau est `VERY_HIGH` confidentielle, elle devient la « première réponse » et est surlignée. Pour les valeurs de cellule de tableau qui sont `HIGH` fiables, elles sont mises en surbrillance. Pour les valeurs de cellule de tableau qui sont `MEDIUM` confidentielles, elles ne sont pas surlignées. La confiance globale pour la réponse du tableau est renvoyée dans la réponse. Par exemple, si un tableau contient principalement des cellules fiables, le niveau de `HIGH` confiance global renvoyé dans la réponse pour la réponse du tableau est le niveau de `HIGH` confiance.

Par défaut, les tableaux n'ont pas une importance ou un poids supérieurs à ceux des autres composants d'un document. Dans un document, si un tableau n'est que légèrement pertinent pour une requête, mais qu'il contient un paragraphe très pertinent, Amazon Kendra renvoie un extrait du paragraphe. Les résultats de recherche affichent le contenu qui fournit la meilleure réponse possible et les informations les plus utiles, dans le même document ou dans d'autres documents. Si le niveau de confiance d'un tableau est inférieur à celui de `MEDIUM` confiance, l'extrait du tableau n'est pas renvoyé dans la réponse.

Pour utiliser la recherche tabulaire sur un index existant, vous devez réindexer votre contenu.

Amazon Kendra la recherche tabulaire prend en charge les synonymes (y compris les synonymes personnalisés). Amazon Kendra ne prend en charge que les documents en anglais contenant des tableaux HTML inclus dans la balise table.

L'exemple suivant montre un extrait de table inclus dans le résultat de la requête. Pour afficher un exemple de JSON avec des réponses à des requêtes, y compris des extraits de tableau, voir Réponses et types de requêtes.

Python

```
import boto3
import pprint

kendra = boto3.client("kendra")

# Provide the index ID
index_id = <index-id>
# Provide the query text
query = "search string"

response = kendra.query(
        QueryText = query,
        IndexId = index_id)

print("\nSearch results for query: " + query + "\n")

for query_result in response["ResultItems"]:

    print("-------------------")
    print("Type: " + str(query_result["Type"]))
    print("Type: " + str(query_result["Format"]))

    if query_result["Type"]=="ANSWER" and query_result["Format"]=="TABLE":
        answer_table = query_result["TableExcerpt"]
        print(answer_table)

    if query_result["Type"]=="ANSWER" and query_result["Format"]=="TEXT":
        answer_text = query_result["DocumentExcerpt"]
        print(answer_text)

    if query_result["Type"]=="QUESTION_ANSWER":
```

```
        question_answer_text = query_result["DocumentExcerpt"]["Text"]
        print(question_answer_text)

    if query_result["Type"]=="DOCUMENT":
        if "DocumentTitle" in query_result:
            document_title = query_result["DocumentTitle"]["Text"]
            print("Title: " + document_title)
        document_text = query_result["DocumentExcerpt"]["Text"]
        print(document_text)

    print("------------------\n\n")
```

Java

```
package com.amazonaws.kendra;

import software.amazon.awssdk.services.kendra.KendraClient;
import software.amazon.awssdk.services.kendra.model.QueryRequest;
import software.amazon.awssdk.services.kendra.model.QueryResponse;
import software.amazon.awssdk.services.kendra.model.QueryResultItem;

public class SearchIndexExample {
    public static void main(String[] args) {
        KendraClient kendra = KendraClient.builder().build();

        String query = "search string";
        String indexId = "index-id";

        QueryRequest queryRequest = QueryRequest
            .builder()
            .queryText(query)
            .indexId(indexId)
            .build();

        QueryResponse queryResponse = kendra.query(queryRequest);

        System.out.println(String.format("\nSearch results for query: %s", query));
        for(QueryResultItem item: queryResponse.resultItems()) {
            System.out.println("----------------------");
            System.out.println(String.format("Type: %s", item.type()));
            System.out.println(String.format("Format: %s", item.format()));

            switch(item.format()) {
```

```
                case TABLE:
                    String answerTable = item.TableExcerpt();
                    System.out.println(answerTable);
                    break;
            }

            switch(item.format()) {
                case TEXT:
                    String answerText = item.DocumentExcerpt();
                    System.out.println(answerText);
                    break;
            }

            switch(item.type()) {
                case QUESTION_ANSWER:
                    String questionAnswerText = item.documentExcerpt().text();
                    System.out.println(questionAnswerText);
                    break;
                case DOCUMENT:
                    String documentTitle = item.documentTitle().text();
                    System.out.println(String.format("Title: %s", documentTitle));
                    String documentExcerpt = item.documentExcerpt().text();
                    System.out.println(String.format("Excerpt: %s",
 documentExcerpt));
                    break;
                default:
                    System.out.println(String.format("Unknown query result type:
 %s", item.type()));

            }

            System.out.println("----------------------\n");
        }
    }
}
```

suggestions de requêtes

> **Note**
>
> La prise en charge des fonctionnalités varie en fonction du type d'index et de l'API de recherche utilisés. Pour savoir si cette fonctionnalité est prise en charge pour le type d'index et l'API de recherche que vous utilisez, consultez la section Types d'index.

Amazon Kendra Les suggestions de requêtes peuvent aider vos utilisateurs à saisir leurs requêtes de recherche plus rapidement et à orienter leur recherche.

Amazon Kendra suggère des requêtes pertinentes pour vos utilisateurs sur la base de l'un des éléments suivants :

- Requêtes populaires dans l'historique des requêtes ou le journal des requêtes
- Le contenu des champs/attributs du document

Vous pouvez définir vos préférences pour l'utilisation de l'historique des requêtes ou des champs du document en les définissant `SuggestionTypes` comme `QUERY` suit ou `DOCUMENT_ATTRIBUTES` en appelant GetQuerySuggestions. Amazon Kendra Utilise par défaut l'historique des requêtes pour baser ses suggestions. Si l'historique des requêtes et les champs du document sont tous deux activés lorsque vous appelez UpdateQuerySuggestionsConfiget vous n'avez pas défini votre `SuggestionTypes` préférence pour utiliser les champs du document, puis Amazon Kendra utilisez l'historique des requêtes.

Si vous utilisez la console, vous pouvez baser vos suggestions de requêtes sur l'historique des requêtes ou sur les champs du document. Vous devez d'abord sélectionner votre index, puis sélectionner Suggestions de requêtes sous Enrichissements dans le menu de navigation. Sélectionnez ensuite Configurer les suggestions de requêtes. Après avoir configuré les suggestions de requêtes, vous êtes dirigé vers une console de recherche dans laquelle vous pouvez sélectionner les champs Historique des requêtes ou Document dans le panneau de droite et saisir une requête de recherche dans la barre de recherche.

Par défaut, les suggestions de requêtes utilisant l'historique des requêtes et les champs du document sont toutes deux activées sans frais supplémentaires. Vous pouvez désactiver ces types de suggestions de requêtes à tout moment à l'aide de l'`UpdateQuerySuggestionsConfig`API. Pour désactiver les suggestions de requêtes en fonction de l'historique des requêtes, définissez

cette option sur `DISABLED` lors `Mode` de l'appel`UpdateQuerySuggestionsConfig`. Pour désactiver les suggestions de requêtes basées sur les champs du document, définissez `AttributeSuggestionsMode` ce paramètre `INACTIVE` dans la configuration des champs du document, puis appelez`UpdateQuerySuggestionsConfig>`. Si vous utilisez la console, vous pouvez désactiver les suggestions de requêtes dans les paramètres des suggestions de requêtes.

Les suggestions de requêtes ne distinguent pas les majuscules et minuscules. Amazon Kendra convertit le préfixe de requête et la requête suggérée en minuscules, ignore tous les guillemets simples et doubles et remplace plusieurs espaces blancs par un seul espace. Amazon Kendra correspond à tous les autres caractères spéciaux tels quels. Amazon Kendra n'affiche aucune suggestion si un utilisateur saisit moins de deux caractères ou plus de 60 caractères.

Rubriques

- Suggestions de requêtes utilisant l'historique des requêtes
- Suggestions de requêtes à l'aide de champs
- Empêcher les suggestions de certaines requêtes ou de certains contenus de champs de documents

Suggestions de requêtes utilisant l'historique des requêtes

> **Note**
>
> La prise en charge des fonctionnalités varie en fonction du type d'index et de l'API de recherche utilisés. Pour savoir si cette fonctionnalité est prise en charge pour le type d'index et l'API de recherche que vous utilisez, consultez la section Types d'index.

Rubriques

- Paramètres de sélection des requêtes pour les suggestions
- Suggestions claires tout en conservant l'historique des requêtes
- Aucune suggestion disponible

Vous pouvez choisir de suggérer des requêtes pertinentes pour vos utilisateurs en fonction des requêtes courantes figurant dans l'historique des requêtes ou dans le journal des requêtes. Amazon Kendra utilise toutes les requêtes recherchées par vos utilisateurs et en tire des leçons pour faire des suggestions à vos utilisateurs. Amazon Kendra suggère des requêtes populaires aux utilisateurs

lorsqu'ils commencent à taper leur requête. Amazon Kendra suggère une requête si le préfixe ou les premiers caractères de la requête correspondent à ce que l'utilisateur commence à taper comme requête.

Par exemple, un utilisateur commence à taper la requête « événements à venir ». Amazon Kendra a appris de l'historique des requêtes que de nombreux utilisateurs ont recherché à de nombreuses reprises « événements à venir en 2050 ». L'utilisateur voit « événements à venir 2050 » apparaître directement sous sa barre de recherche, complétant automatiquement sa requête de recherche. L'utilisateur sélectionne cette suggestion de requête et le document « Nouveaux événements : que se passe-t-il en 2050 » apparaît dans les résultats de recherche.

Vous pouvez définir le mode Amazon Kendra de sélection des requêtes éligibles à suggérer à vos utilisateurs. Par exemple, vous pouvez spécifier qu'une suggestion de requête doit avoir été recherchée par au moins 10 utilisateurs uniques (la valeur par défaut est trois), avoir été recherchée au cours des 30 derniers jours et ne contenir aucun mot ou expression de votre liste de blocage. Amazon Kendra exige qu'une requête comporte au moins un résultat de recherche et contienne au moins un mot de plus de quatre caractères.

Paramètres de sélection des requêtes pour les suggestions

Vous pouvez configurer les paramètres suivants pour sélectionner des requêtes de suggestions à l'aide du UpdateQuerySuggestionsConfigAPI :

- Mode —Les suggestions de requêtes utilisant l'historique des requêtes sont soit, `ENABLED` soit`LEARN_ONLY`. Amazon Kendra active les suggestions de requêtes par défaut. `LEARN_ONLY`désactive les suggestions de requêtes. Si cette option est désactivée, elle Amazon Kendra continue à apprendre les suggestions, mais ne propose pas de suggestions de requêtes aux utilisateurs.
- Fenêtre temporelle du journal des requêtes : date de la date de vos requêtes dans la fenêtre temporelle du journal des requêtes. La fenêtre temporelle est une valeur entière correspondant au nombre de jours entre le jour actuel et les jours précédents.
- Requêtes sans informations utilisateur : définissez ce `TRUE` paramètre sur pour inclure toutes les requêtes ou sur uniquement `FALSE` les requêtes contenant des informations utilisateur. Vous pouvez utiliser ce paramètre si votre application de recherche inclut des informations utilisateur, telles que l'ID utilisateur, lorsqu'un utilisateur émet une requête. Par défaut, ce paramètre ne filtre pas les requêtes si aucune information utilisateur spécifique n'est associée aux requêtes. Toutefois, vous ne pouvez utiliser ce paramètre que pour faire des suggestions basées sur des requêtes contenant des informations utilisateur.

- Utilisateurs uniques : nombre minimum d'utilisateurs uniques qui doivent effectuer une recherche pour que celle-ci puisse être suggérée à vos utilisateurs. Ce nombre est une valeur entière.
- Nombre de requêtes : nombre minimum de fois qu'une requête doit être recherchée pour qu'elle puisse être suggérée à vos utilisateurs. Ce nombre est une valeur entière.

Ces paramètres affectent la manière dont les requêtes sont sélectionnées en tant que requêtes populaires à suggérer à vos utilisateurs. La façon dont vous réglerez vos paramètres dépendra de vos besoins spécifiques, par exemple :

- Si vos utilisateurs effectuent généralement une recherche une fois par mois en moyenne, vous pouvez définir le nombre de jours dans la fenêtre du journal des requêtes à 30 jours. En utilisant ce paramètre, vous capturez la plupart des requêtes récentes de vos utilisateurs avant qu'elles ne deviennent obsolètes dans la fenêtre temporelle.
- Si seul un petit nombre de vos requêtes incluent des informations utilisateur et que vous ne souhaitez pas suggérer de requêtes basées sur un échantillon de petite taille, vous pouvez définir des requêtes pour inclure tous les utilisateurs.
- Si vous définissez les requêtes populaires comme étant recherchées par au moins 10 utilisateurs uniques et recherchées au moins 100 fois, vous définissez les utilisateurs uniques sur 10 et le nombre de requêtes sur 100.

> **Warning**
>
> Les modifications que vous avez apportées aux paramètres peuvent ne pas prendre effet immédiatement. Vous pouvez suivre les modifications des paramètres en utilisant le DescribeQuerySuggestionsConfigAPI. Le délai d'entrée en vigueur de vos paramètres mis à jour dépend des mises à jour que vous effectuez et du nombre de requêtes de recherche dans votre index. Amazon Kendra met automatiquement à jour les suggestions toutes les 24 heures, après avoir modifié un paramètre ou après avoir appliqué une liste de blocage.

CLI

Pour récupérer des suggestions de requêtes

```
aws kendra get-query-suggestions \
 --index-id index-id \
 --query-text "query-text" \
```

```
 --suggestion-types '["QUERY"]' \
 --max-suggestions-count 1 // If you want to limit the number of suggestions
```

Pour mettre à jour les suggestions de requêtes

Par exemple, pour modifier la fenêtre temporelle du journal des requêtes et le nombre minimum de fois qu'une requête doit être recherchée :

```
aws kendra update-query-suggestions-config \
 --index-id index-id \
 --query-log-look-back-window-in-days 30 \
 --minimum-query-count 100
```

Python

Pour récupérer des suggestions de requêtes

```
import boto3
from botocore.exceptions import ClientError

kendra = boto3.client("kendra")

print("Get query suggestions.")

# Provide the index ID
index_id = "index-id"

# Provide the query text
query_text = "query"

# Provide the query suggestions type
query_suggestions_type = "QUERY"


# If you want to limit the number of suggestions
num_suggestions = 1

try:
    query_suggestions_response = kendra.get_query_suggestions(
        IndexId = index_id,
        QueryText = query_text,
        SuggestionTypes = query_suggestions_type,
        MaxSuggestionsCount = num_suggestions
```

```
    )

    # Print out the suggestions you received
    if ("Suggestions" in query_suggestions_response.keys()) {
        for (suggestion: query_suggestions_response["Suggestions"]) {
            print(suggestion["Value"]["Text"]["Text"]);
        }
    }

except ClientError as e:
        print("%s" % e)

print("Program ends.")
```

Pour mettre à jour les suggestions de requêtes

Par exemple, pour modifier la fenêtre temporelle du journal des requêtes et le nombre minimum de fois qu'une requête doit être recherchée :

```
import boto3
from botocore.exceptions import ClientError
import pprint
import time

kendra = boto3.client("kendra")

print("Updating query suggestions settings/configuration for an index.")

# Provide the index ID
index_id = "index-id"

# Configure the settings you want to update
minimum_query_count = 100
query_log_look_back_window_in_days = 30

try:
    kendra.update_query_suggestions_config(
        IndexId = index_id,
        MinimumQueryCount = minimum_query_count,
        QueryLogLookBackWindowInDays = query_log_look_back_window_in_days
    )

    print("Wait for Amazon Kendra to update the query suggestions.")
```

```
    while True:
        # Get query suggestions description of settings/configuration
        query_sugg_config_response = kendra.describe_query_suggestions_config(
            IndexId = index_id
        )

        # If status is not UPDATING, then quit
        status = query_sugg_config_response["Status"]
        print(" Updating query suggestions config. Status: " + status)
        if status != "UPDATING":
            break
        time.sleep(60)

except ClientError as e:
        print("%s" % e)

print("Program ends.")
```

Suggestions claires tout en conservant l'historique des requêtes

> **Note**
>
> La prise en charge des fonctionnalités varie en fonction du type d'index et de l'API de recherche utilisés. Pour savoir si cette fonctionnalité est prise en charge pour le type d'index et l'API de recherche que vous utilisez, consultez la section Types d'index.

Vous pouvez effacer les suggestions de requêtes en utilisant le ClearQuerySuggestionsAPI. La suppression des suggestions supprime uniquement les suggestions de requêtes existantes, et non les requêtes de l'historique des requêtes. Lorsque vous effacez des suggestions, Amazon Kendra apprend les nouvelles suggestions en fonction des nouvelles requêtes ajoutées au journal des requêtes à partir du moment où vous avez effacé les suggestions.

CLI

Pour effacer les suggestions de requêtes

```
aws  kendra clear-query-suggestions \
 --index-id index-id
```

Python

Pour effacer les suggestions de requêtes

```
import boto3
from botocore.exceptions import ClientError

kendra = boto3.client("kendra")

print("Clearing out query suggestions for an index.")

# Provide the index ID
index_id = "index-id"

try:
    kendra.clear_query_suggestions(
        IndexId = index_id
    )

    # Confirm last cleared date-time and that there are no suggestions
    query_sugg_config_response = kendra.describe_query_suggestions_config(
        IndexId = index_id
    )
    print("Query Suggestions last cleared at: " +
 str(query_sugg_config_response["LastClearTime"]));
    print("Number of suggestions available from the time of clearing: " +
 str(query_sugg_config_response["TotalSuggestionsCount"]));

except ClientError as e:
        print("%s" % e)

print("Program ends.")
```

Aucune suggestion disponible

Si aucune suggestion de requête ne s'affiche, cela peut être dû à l'une des raisons suivantes :

- Votre index ne contient pas suffisamment de requêtes pour en Amazon Kendra tirer des leçons.
- Vos paramètres de suggestions de requêtes sont trop stricts, de sorte que la plupart des requêtes sont exclues des suggestions.

- Vous avez récemment effacé des suggestions et vous avez Amazon Kendra encore besoin de temps pour que les nouvelles requêtes s'accumulent afin de connaître de nouvelles suggestions.

Vous pouvez vérifier vos paramètres actuels à l'aide du DescribeQuerySuggestionsConfigAPI.

Suggestions de requêtes à l'aide de champs

Rubriques

- Paramètres de sélection des champs pour les suggestions
- Contrôle par l'utilisateur dans les champs du document

Vous pouvez choisir de suggérer des requêtes pertinentes pour vos utilisateurs en fonction du contenu des champs du document. Au lieu d'utiliser l'historique des requêtes pour suggérer d'autres requêtes pertinentes populaires, vous pouvez utiliser les informations contenues dans un champ de document qui sont utiles pour compléter automatiquement la requête. Amazon Kendra recherche le contenu pertinent dans les champs définis sur `Suggestable` et correspondant étroitement à la requête de votre utilisateur. Amazon Kendra Suggère ensuite ce contenu à votre utilisateur lorsqu'il commence à taper sa requête.

Par exemple, si vous spécifiez le champ de titre sur lequel baser les suggestions et qu'un utilisateur commence à taper la requête « How amazon ken... », le titre le plus pertinent « Comment Amazon Kendra fonctionne » pourrait être suggéré pour compléter automatiquement la recherche. L'utilisateur voit « Comment Amazon Kendra fonctionne » apparaître directement sous sa barre de recherche, complétant automatiquement sa requête de recherche. L'utilisateur sélectionne cette suggestion de requête, et le document « Comment Amazon Kendra fonctionne » apparaît dans les résultats de recherche.

Vous pouvez utiliser le contenu de n'importe quel champ `String` et `StringList` type de document pour suggérer une requête en définissant le champ sur dans le cadre de `Suggestable` la configuration de vos champs pour les suggestions de requêtes. Vous pouvez également utiliser une liste de blocage afin que les champs de document suggérés contenant certains mots ou expressions ne soient pas affichés à vos utilisateurs. Vous pouvez utiliser une seule liste de blocage. La liste de blocage s'applique, que vous définissiez des suggestions de requêtes pour utiliser l'historique des requêtes ou les champs du document.

Paramètres de sélection des champs pour les suggestions

Vous pouvez configurer les paramètres suivants pour sélectionner les champs du document pour les suggestions à l'aide de AttributeSuggestionsConfiget en appelant le UpdateQuerySuggestionsConfigAPI pour mettre à jour les paramètres au niveau de l'index :

- Mode suggestions de champs/attributs : les suggestions de requêtes utilisant des champs de document sont soit, soit. `ACTIVE` `INACTIVE` Amazon Kendra active les suggestions de requêtes par défaut.
- Champs/attributs suggestibles : noms de champs ou clés de champ sur lesquels baser les suggestions. Ces champs doivent être définis sur `TRUE` for`Suggestable`, dans le cadre de la configuration des champs. Vous pouvez modifier la configuration des champs au niveau de la requête tout en conservant la configuration au niveau de l'index. Utilisez l'GetQuerySuggestionsAPI pour modifier `AttributeSuggestionConfig` au niveau de la requête. Cette configuration au niveau de la requête peut être utile pour expérimenter rapidement l'utilisation de différents champs de document sans avoir à mettre à jour la configuration au niveau de l'index.
- Champs/attributs supplémentaires : champs supplémentaires que vous souhaitez inclure dans la réponse à une suggestion de requête. Ces champs sont utilisés pour fournir des informations supplémentaires dans la réponse ; toutefois, ils ne sont pas utilisés pour baser les suggestions.

> **Warning**
>
> Les modifications que vous avez apportées aux paramètres peuvent ne pas prendre effet immédiatement. Vous pouvez suivre les modifications des paramètres en utilisant le DescribeQuerySuggestionsConfigAPI. Le délai d'entrée en vigueur de vos paramètres mis à jour dépend des mises à jour que vous effectuez. Amazon Kendra met automatiquement à jour les suggestions toutes les 24 heures, après avoir modifié un paramètre ou après avoir appliqué une liste de blocage.

CLI

Pour récupérer des suggestions de requêtes et modifier la configuration des champs du document au niveau de la requête au lieu de modifier la configuration au niveau de l'index.

```
aws kendra get-query-suggestions \
 --index-id index-id \
 --query-text "query-text" \
```

```
 --suggestion-types '["DOCUMENT_ATTRIBUTES"]' \
 --attribute-suggestions-config '{"SuggestionAttributes":'["field/attribute key
 1", "field/attribute key 2"]', "AdditionalResponseAttributes":'["response field/
attribute key 1", "response field/attribute key 2"]'}' \
 --max-suggestions-count 1 // If you want to limit the number of suggestions
```

Pour mettre à jour les suggestions de requêtes

Par exemple, pour modifier la configuration des champs du document au niveau de l'index :

```
aws kendra update-query-suggestions-config \
 --index-id index-id \
 --attribute-suggestions-config '{"SuggestableConfigList": '[{"SuggestableConfig":
 "_document_title", "Suggestable": true}]', "AttributeSuggestionsMode": "ACTIVE"}'
```

Python

Pour récupérer des suggestions de requêtes et modifier la configuration des champs du document au niveau de la requête au lieu de modifier la configuration au niveau de l'index.

```
import boto3
from botocore.exceptions import ClientError

kendra = boto3.client("kendra")

print("Get query suggestions.")

# Provide the index ID
index_id = "index-id"

# Provide the query text
query_text = "query"

# Provide the query suggestions type
query_suggestions_type = "DOCUMENT_ATTRIBUTES"

# Override fields/attributes configuration at query level
configuration = {"SuggestionAttributes":
    '["field/attribute key 1", "field/attribute key 2"]',
      "AdditionalResponseAttributes":
          '["response field/attribute key 1", "response field/attribute key 2"]'
          }
```

```
# If you want to limit the number of suggestions
num_suggestions = 1

try:
    query_suggestions_response = kendra.get_query_suggestions(
        IndexId = index_id,
        QueryText = query_text,
        SuggestionTypes = [query_suggestions_type],
        AttributeSuggestionsConfig = configuration,
        MaxSuggestionsCount = num_suggestions
    )

    # Print out the suggestions you received
    if ("Suggestions" in query_suggestions_response.keys()) {
        for (suggestion: query_suggestions_response["Suggestions"]) {
            print(suggestion["Value"]["Text"]["Text"]);
        }
    }

except ClientError as e:
        print("%s" % e)

print("Program ends.")
```

Pour mettre à jour les suggestions de requêtes

Par exemple, pour modifier la configuration des champs du document au niveau de l'index :

```
import boto3
from botocore.exceptions import ClientError
import pprint
import time

kendra = boto3.client("kendra")

print("Updating query suggestions settings/configuration for an index.")

# Provide the index ID
index_id = "index-id"

# Configure the settings you want to update at the index level
configuration = {"SuggestableConfigList":
    '[{"SuggestableConfig": "_document_title", "Suggestable": true}]',
       "AttributeSuggestionsMode": "ACTIVE"
```

```
        }

try:
    kendra.update_query_suggestions_config(
        IndexId = index_id,
        AttributeSuggestionsConfig = configuration
    )

    print("Wait for Amazon Kendra to update the query suggestions.")

    while True:
        # Get query suggestions description of settings/configuration
        query_sugg_config_response = kendra.describe_query_suggestions_config(
            IndexId = index_id
        )

        # If status is not UPDATING, then quit
        status = query_sugg_config_response["Status"]
        print(" Updating query suggestions config. Status: " + status)
        if status != "UPDATING":
            break
        time.sleep(60)

except ClientError as e:
        print("%s" % e)

print("Program ends.")
```

Contrôle par l'utilisateur dans les champs du document

Vous pouvez appliquer un filtrage contextuel utilisateur aux champs du document sur lesquels vous souhaitez baser les suggestions de requêtes. Cela filtre les informations des champs du document en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Par exemple, un stagiaire effectue une recherche sur le portail de l'entreprise et n'a pas accès à un document top secret de l'entreprise. Par conséquent, les requêtes suggérées basées sur le titre du document top secret, ou sur tout autre champ suggestible, ne sont pas présentées au stagiaire.

Vous pouvez indexer vos documents à l'aide d'une liste de contrôle d'accès (ACL), qui définit quels utilisateurs et groupes sont autorisés à accéder à quels documents. Vous pouvez ensuite appliquer un filtrage contextuel utilisateur aux champs de vos documents pour les suggestions de requêtes. Le filtrage du contexte utilisateur actuellement défini pour votre index est le même que celui appliqué

à la configuration des champs de votre document pour les suggestions de requêtes. Le filtrage du contexte utilisateur fait partie de la configuration des champs de votre document. Vous utilisez le AttributeSuggestionsGetConfiget appelez GetQuerySuggestions.

Empêcher les suggestions de certaines requêtes ou de certains contenus de champs de documents

Une liste de blocage Amazon Kendra empêche de suggérer certaines requêtes à vos utilisateurs. Une liste de blocage est une liste de mots ou d'expressions que vous souhaitez exclure des suggestions de requêtes. Amazon Kendra exclut les requêtes contenant une correspondance exacte avec les mots ou les phrases de la liste de blocage.

Vous pouvez utiliser une liste de blocage pour vous protéger contre les mots ou expressions offensants qui apparaissent fréquemment dans l'historique de vos requêtes ou dans les champs du document et qui Amazon Kendra pourraient être sélectionnés comme suggestions. Une liste de blocage peut également Amazon Kendra empêcher de suggérer des requêtes contenant des informations qui ne sont pas prêtes à être publiées ou annoncées. Par exemple, vos utilisateurs posent fréquemment des questions sur la sortie prochaine d'un nouveau produit potentiel. Toutefois, vous ne souhaitez pas suggérer le produit car vous n'êtes pas prêt à le lancer. Vous pouvez bloquer les requêtes contenant le nom du produit et des informations sur le produit dans les suggestions.

Vous pouvez créer une liste de blocage pour les requêtes à l'aide du CreateQuerySuggestionsBlockListAPI. Vous placez chaque mot ou phrase en bloc sur une ligne distincte d'un fichier texte. Ensuite, vous chargez le fichier texte dans votre compartiment Amazon S3 et vous indiquez le chemin ou l'emplacement du fichier Amazon S3. Amazon Kendra prend actuellement en charge la création d'une seule liste de blocage.

Vous pouvez remplacer le fichier texte contenant les mots et expressions bloqués dans votre Amazon S3 compartiment. Pour mettre à jour la liste de blocage dans Amazon Kendra, utilisez le UpdateQuerySuggestionsBlockListAPI.

Utilisation de la DescribeQuerySuggestionsBlockListAPI pour obtenir le statut de votre liste de blocage. `DescribeQuerySuggestionsBlockList`peut également vous fournir d'autres informations utiles, telles que les suivantes :

- Quand votre liste de blocage a été mise à jour pour la dernière fois
- Combien de mots ou d'expressions se trouvent dans votre liste de blocage actuelle
- Messages d'erreur utiles lors de la création d'une liste de blocage

Vous pouvez également utiliser le ListQuerySuggestionsBlockListsAPI pour obtenir une liste de résumés de listes de blocage pour un index.

Pour supprimer votre liste de blocage, utilisez l'DeleteQuerySuggestionsBlockListAPI.

Les mises à jour de la liste de blocage risquent de ne pas prendre effet immédiatement. Vous pouvez suivre les mises à jour à l'aide de l'`DescribeQuerySuggestionsBlockList`API.

CLI

Pour créer une liste de blocage

```
aws kendra create-query-suggestions-block-list \
 --index-id index-id \
 --name "block-list-name" \
 --description "block-list-description" \
 --source-s3-path "Bucket=bucket-name,Key=query-suggestions/block_list.txt" \
 --role-arn role-arn
```

Pour mettre à jour une liste de blocage

```
aws kendra update-query-suggestions-block-list \
 --index-id index-id \
 --name "new-block-list-name" \
 --description "new-block-list-description" \
 --source-s3-path "Bucket=bucket-name,Key=query-suggestions/new_block_list.txt" \
 --role-arn role-arn
```

Pour supprimer une liste de blocage

```
aws kendra delete-query-suggestions-block-list \
 --index-id index-id \
 --id block-list-id
```

Python

Pour créer une liste de blocage

```
import boto3
from botocore.exceptions import ClientError
import pprint
import time
```

```
kendra = boto3.client("kendra")

print("Create a query suggestions block list.")

# Provide a name for the block list
block_list_name = "block-list-name"
# Provide an optional description for the block list
block_list_description = "block-list-description"
# Provide the IAM role ARN required for query suggestions block lists
block_list_role_arn = "role-arn"

# Provide the index ID
index_id = "index-id"

s3_bucket_name = "bucket-name"
s3_key = "query-suggestions/block_list.txt"
source_s3_path = {
    'Bucket': s3_bucket_name,
    'Key': s3_key
}

try:
    block_list_response = kendra.create_query_suggestions_block_list(
        Description = block_list_description,
        Name = block_list_name,
        RoleArn = block_list_role_arn,
        IndexId = index_id,
        SourceS3Path = source_s3_path
    )

    print(block_list_response)

    block_list_id = block_list_response["Id"]

    print("Wait for Amazon Kendra to create the block list.")

    while True:
        # Get block list description
        block_list_description = kendra.describe_query_suggestions_block_list(
            Id = block_list_id,
            IndexId = index_id
        )
        # If status is not CREATING, then quit
```

```
        status = block_list_description["Status"]
        print("Creating block list. Status: " + status)
        if status != "CREATING":
            break
        time.sleep(60)

except ClientError as e:
        print("%s" % e)

print("Program ends.")
```

Pour mettre à jour une liste de blocage

```
import boto3
from botocore.exceptions import ClientError
import pprint
import time

kendra = boto3.client("kendra")

print("Update a block list for query suggestions.")

# Provide the block list name you want to update
block_list_name = "new-block-list-name"
# Provide the block list description you want to update
block_list_description = "new-block-list-description"
# Provide the IAM role ARN required for query suggestions block lists
block_list_role_arn = "role-arn"

# Provide the block list ID
block_list_id = "block-list-id"
# Provide the index ID
index_id = "index-id"

s3_bucket_name = "bucket-name"
s3_key = "query-suggestions/new_block_list.txt"
source_s3_path = {
'Bucket': s3_bucket_name,
'Key': s3_key
}

try:
    kendra.update_query_suggestions_block_list(
```

```
        Id = block_list_id,
        IndexId = index_id,
        Description = block_list_description,
        Name = block_list_name,
        RoleArn = block_list_role_arn,
        SourceS3Path = source_s3_path
    )

    print("Wait for Amazon Kendra to update the block list.")

    while True:
        # Get block list description
        block_list_description = kendra.describe_query_suggestions_block_list(
            Id = block_list_id,
            IndexId = index_id
        )
        # If status is not UPDATING, then the update has finished
        status = block_list_description["Status"]
        print("Updating block list. Status: " + status)
        if status != "UPDATING":
            break
        time.sleep(60)

except ClientError as e:
print("%s" % e)

print("Program ends.")
```

Pour supprimer une liste de blocage

```
import boto3
from botocore.exceptions import ClientError

kendra = boto3.client("kendra")

print("Delete a block list for query suggestions.")

# provide the block list ID
query_suggestions_block_list_id = "query-suggestions-block-list-id"
# Provide the index ID
index_id = "index-id"

try:
```

```
    kendra.delete_query_suggestions_block_list(
        Id = query_suggestions_block_list_id,
        IndexId = index_id
    )

except ClientError as e:
        print("%s" % e)

print("Program ends.")
```

Correcteur orthographique des requêtes

> **Note**
>
> La prise en charge des fonctionnalités varie en fonction du type d'index et de l'API de recherche utilisés. Pour savoir si cette fonctionnalité est prise en charge pour le type d'index et l'API de recherche que vous utilisez, consultez la section Types d'index.

Amazon Kendra Le correcteur orthographique suggère des corrections orthographiques pour une requête. Cela peut vous aider à réduire au minimum les occurrences de zéro résultat de recherche et à renvoyer des résultats pertinents. Vos utilisateurs peuvent ne recevoir aucun résultat de recherche en raison de requêtes mal orthographiées sans résultats correspondants ou sans documents renvoyés. Il se peut également que vos utilisateurs reçoivent des résultats de recherche non pertinents en raison de requêtes mal orthographiées.

Le correcteur orthographique est conçu pour suggérer des corrections pour les mots mal orthographiés en fonction des mots qui apparaissent dans vos documents indexés et de la mesure dans laquelle un mot corrigé correspond à un mot mal orthographié. Par exemple, si le mot « relevés » apparaît dans vos documents indexés, il peut correspondre étroitement au mot mal orthographié « relevés » dans la requête « états financiers de fin d'année ».

Le correcteur orthographique renvoie les mots prévus ou corrigés qui remplacent les mots mal orthographiés dans le texte de la requête d'origine. Par exemple, « déployer la recherche Kendre » peut renvoyer « déployer la recherche Kendra ». Vous pouvez également utiliser les emplacements de décalage fournis dans l'API pour surligner ou mettre en italique les mots corrigés renvoyés dans une requête de votre application frontale. Dans la console, les mots corrigés sont surlignés ou en italique par défaut. Par exemple, « déployer la recherche Kendra ».

Pour les termes spécifiques à une entreprise ou spécialisés qui apparaissent dans vos documents indexés, le correcteur orthographique ne les interprète pas à tort comme des fautes d'orthographe dans la requête. Par exemple, « Amazon Macie » n'est pas remplacé par « Amazon Mace ».

Pour les mots comportant un trait d'union, tels que « fin d'année », le correcteur orthographique les traite comme des mots individuels afin de suggérer des corrections pour ces mots. Par exemple, la correction suggérée pour « fin d'année » pourrait être « fin d'année ».

Pour les types de réponses aux `QUESTION_ANSWER` requêtes `DOCUMENT` et aux requêtes, le correcteur orthographique suggère de corriger les mots mal orthographiés en fonction des mots présents dans le corps du document. Le corps du document est plus fiable que le titre pour suggérer des corrections correspondant étroitement aux mots mal orthographiés. Pour les types de réponses aux `ANSWER` requêtes, le correcteur orthographique suggère des corrections basées sur les mots figurant dans le document de questions-réponses par défaut de votre index.

Vous pouvez activer le correcteur orthographique à l'aide de l'SpellCorrectionConfigurationobjet. Vous avez réglé `IncludeQuerySpellCheckSuggestions` sur`TRUE`. Le correcteur orthographique est activé par défaut dans la console. Il est intégré à la console par défaut.

Le correcteur orthographique peut également suggérer des corrections orthographiques pour les requêtes dans plusieurs langues, pas seulement en anglais. Pour obtenir la liste des langues prises en charge par le correcteur orthographique, consultez la section Langues Amazon Kendra prises en charge.

Utilisation du correcteur orthographique des requêtes avec des limites par défaut

Le correcteur orthographique est conçu avec certaines valeurs par défaut ou limites. Voici une liste des limites actuelles qui s'appliquent lorsque vous activez les suggestions de correction orthographique.

- Les corrections orthographiques suggérées ne peuvent pas être renvoyées pour les mots de moins de trois caractères ou de plus de 30 caractères. Pour autoriser plus de 30 caractères ou moins de trois caractères, contactez le Support.
- Les corrections orthographiques suggérées ne peuvent pas restreindre les suggestions basées sur le contrôle d'accès des utilisateurs ou sur votre liste de contrôle d'accès pour le filtrage du contexte utilisateur. Les corrections orthographiques sont basées sur tous les mots de vos documents indexés, qu'ils soient réservés à certains utilisateurs ou non. Si vous souhaitez éviter que certains

mots apparaissent dans les corrections orthographiques suggérées pour les requêtes, n'activez pas`SpellCorrectionConfiguration`.

- Les corrections orthographiques suggérées ne peuvent pas être renvoyées pour les mots contenant des chiffres. Par exemple, « how 2 not br8k ubun2 ».
- Les corrections orthographiques suggérées ne peuvent pas utiliser de mots qui n'apparaissent pas dans vos documents indexés.
- Les corrections orthographiques suggérées ne peuvent pas utiliser de mots qui sont fréquentés à moins de 0,01 % dans vos documents indexés. Pour modifier le seuil de 0,01 %, contactez le Support.

Filtrage et recherche par facettes

> **Note**
>
> La prise en charge des fonctionnalités varie en fonction du type d'index et de l'API de recherche utilisés. Pour savoir si cette fonctionnalité est prise en charge pour le type d'index et l'API de recherche que vous utilisez, consultez la section Types d'index.

Vous pouvez améliorer les résultats de recherche ou les réponses de l'API Query en utilisant des filtres. Les filtres limitent les documents de la réponse à ceux qui s'appliquent directement à la requête. Pour créer des suggestions de recherche à facettes, utilisez la logique booléenne pour filtrer les attributs de document spécifiques de la réponse ou les documents qui ne correspondent pas à des critères spécifiques. Vous pouvez spécifier des facettes à l'aide du `Facets` paramètre de l'`Query`API.

Pour rechercher des documents que vous avez indexés avec Amazon Kendra for Amazon Lex, utilisez AMAZON. KendraSearchIntent. Pour un exemple de configuration Amazon Kendra avec Amazon Lex, voir Création d'un bot FAQ pour un Amazon Kendra index. Vous pouvez également fournir un filtre pour la réponse en utilisant AttributeFilter. Il s'agit du filtre de requête en JSON lors de la configuration`AMAZON.KendraSearchIntent`. Pour fournir un filtre d'attributs lors de la configuration d'une intention de recherche dans la console, accédez à l'éditeur d'intention et choisissez Amazon Kendra query pour fournir un filtre de requête au format JSON. Pour plus d'informations`AMAZON.KendraSearchIntent`, consultez le guide de Amazon Lex documentation.

Facettes

Les facettes sont des vues étendues d'un ensemble de résultats de recherche. Par exemple, vous pouvez fournir des résultats de recherche pour des villes du monde entier, où les documents sont filtrés en fonction de la ville à laquelle ils sont associés. Vous pouvez également créer des facettes pour afficher les résultats d'un auteur spécifique.

Vous pouvez utiliser un attribut de document ou un champ de métadonnées associé à un document en tant que facette afin que vos utilisateurs puissent effectuer une recherche par catégorie ou par valeur au sein de cette facette. Vous pouvez également afficher des facettes imbriquées dans les résultats de recherche afin que vos utilisateurs puissent effectuer une recherche non seulement par catégorie ou par champ, mais également par sous-catégorie ou sous-champ.

L'exemple suivant montre comment obtenir des informations facettaires pour l'attribut personnalisé « Ville ».

```
response=kendra.query(
        QueryText = query,
        IndexId = index,
        Facets = [
            {
                "DocumentAttributeKey" : "City"
            }
        ]
        )
```

Vous pouvez utiliser des facettes imbriquées pour affiner davantage la recherche. Par exemple, l'attribut ou la facette « Ville » du document inclut une valeur appelée « Seattle ». En outre, l'attribut ou facette du document « CityRegion » inclut les valeurs « Nord » et « Sud » pour les documents affectés à « Seattle ». Vous pouvez afficher les facettes imbriquées avec leur nombre dans les résultats de recherche afin que les documents puissent être recherchés non seulement par ville, mais également par région au sein d'une ville.

Notez que les facettes imbriquées peuvent avoir un impact sur la latence des requêtes. En règle générale, plus vous utilisez de facettes imbriquées, plus l'impact potentiel sur la latence est important. Parmi les autres facteurs qui influent sur la latence, citons la taille moyenne des documents indexés, la taille de votre index, les requêtes très complexes et la charge globale de votre Amazon Kendra index.

L'exemple suivant montre comment obtenir des informations de facette pour l'attribut personnalisé « CityRegion », sous forme de facette imbriquée dans « City ».

```
response=kendra.query(
        QueryText = query,
        IndexId = index,
        Facets = [
            {
                "DocumentAttributeKey" : "City",
                "Facets": [
                    {
                        "DocumentAttributeKey" : "CityRegion"
                    }
                ]
            }
        ]
        )
```

Les informations relatives aux facettes, telles que le nombre de documents, sont renvoyées dans le tableau de `FacetResults` réponses. Vous utilisez le contenu pour afficher des suggestions de recherche à facettes dans votre application. Par exemple, si l'attribut « Ville » du document contient la ville à laquelle une recherche pourrait s'appliquer, utilisez ces informations pour afficher une liste des villes recherchées. Les utilisateurs peuvent choisir une ville pour filtrer leurs résultats de recherche. Pour effectuer la recherche à facettes, appelez l'API Query et utilisez l'attribut de document choisi pour filtrer les résultats.

Vous pouvez afficher jusqu'à 10 valeurs de facette par facette pour une requête, et une seule facette imbriquée par facette. Si vous souhaitez augmenter ces limites, contactez le Support. Si vous souhaitez limiter le nombre de valeurs de facette par facette à moins de 10, vous pouvez le spécifier dans l'`Facet`objet.

L'exemple de réponse JSON suivant montre les facettes limitées à l'attribut de document « Ville ». La réponse inclut le nombre de documents pour la valeur de facette.

```
{
    'FacetResults': [
        {
            'DocumentAttributeKey': 'City',
            'DocumentAttributeValueCountPairs': [
                {
                    'Count': 3,
```

```
                    'DocumentAttributeValue': {
                        'StringValue': 'Dubai'
                    }
                },
                {
                    'Count': 3,
                    'DocumentAttributeValue': {
                        'StringValue': 'Seattle'
                    }
                },
                {
                    'Count': 1,
                    'DocumentAttributeValue': {
                        'StringValue': 'Paris'
                    }
                }
            ]
        }
    ]
```

Vous pouvez également afficher les informations relatives à une facette imbriquée, telle qu'une région au sein d'une ville, afin de filtrer davantage les résultats de recherche.

L'exemple de réponse JSON suivant montre les facettes limitées à l'attribut de document « CityRegion », sous la forme d'une facette imbriquée dans « City ». La réponse inclut le nombre de documents pour les valeurs de facettes imbriquées.

```
{
    'FacetResults': [
        {
            'DocumentAttributeKey': 'City',
            'DocumentAttributeValueCountPairs': [
                {
                    'Count': 3,
                    'DocumentAttributeValue': {
                        'StringValue': 'Dubai'
                    },
                    'FacetResults': [
                        {
                            'DocumentAttributeKey': 'CityRegion',
                            'DocumentAttributeValueCountPairs': [
                                {
                                    'Count': 2,
```

```
                            'DocumentAttributeValue': {
                                'StringValue': 'Bur Dubai'
                            }
                        },
                        {
                            'Count': 1,
                            'DocumentAttributeValue': {
                                'StringValue': 'Deira'
                            }
                        }
                    ]
                }
            ]
        },
        {
            'Count': 3,
            'DocumentAttributeValue': {
                'StringValue': 'Seattle'
            },
            'FacetResults': [
                {
                    'DocumentAttributeKey': 'CityRegion',
                    'DocumentAttributeValueCountPairs': [
                        {
                            'Count': 1,
                            'DocumentAttributeValue': {
                                'StringValue': 'North'
                            }
                        },
                        {
                            'Count': 2,
                            'DocumentAttributeValue': {
                                'StringValue': 'South'
                            }
                        }
                    ]
                }
            ]
        },
        {
            'Count': 1,
            'DocumentAttributeValue': {
                'StringValue': 'Paris'
            },
```

```
                    'FacetResults': [
                        {
                            'DocumentAttributeKey': 'CityRegion',
                            'DocumentAttributeValueCountPairs': [
                                {
                                    'Count': 1,
                                    'DocumentAttributeValue': {
                                        'StringValue': 'City center'
                                    }
                                }
                             ]
                        }
                    ]
                }
        }
    ]
}
```

Lorsque vous utilisez un champ de liste de chaînes pour créer des facettes, les résultats des facettes renvoyés sont basés sur le contenu de la liste de chaînes. Par exemple, si vous avez un champ de liste de chaînes contenant deux éléments, l'un avec la liste « teckel », « saucisson » et l'autre avec la valeur « husky », vous obtenez `FacetResults` trois facettes.

Pour de plus amples informations, veuillez consulter Réponses aux requêtes et types de réponses.

Utilisation des attributs du document pour filtrer les résultats de recherche

Par défaut, `Query` renvoie tous les résultats de recherche. Pour filtrer les réponses, vous pouvez effectuer des opérations logiques sur les attributs du document. Par exemple, si vous souhaitez uniquement des documents pour une ville spécifique, vous pouvez filtrer selon les attributs de document personnalisés « Ville » et « État ». Vous pouvez l'utiliser AttributeFilterpour créer une opération booléenne sur les filtres que vous fournissez.

La plupart des attributs peuvent être utilisés pour filtrer les réponses pour tous les types de réponses. Toutefois, l'`_excerpt_page_number`attribut ne s'applique qu'aux types de `ANSWER` réponses lors du filtrage des réponses.

L'exemple suivant montre comment effectuer une opération logique AND en filtrant sur une ville spécifique, Seattle, et sur l'État, Washington.

```
response=kendra.query(
```

```
        QueryText = query,
        IndexId = index,
        AttributeFilter = {'AndAllFilters':
            [
                {"EqualsTo": {"Key": "City","Value": {"StringValue": "Seattle"}}},
                {"EqualsTo": {"Key": "State","Value": {"StringValue": "Washington"}}}
            ]
            }
        )
```

L'exemple suivant montre comment effectuer une opération logique OR lorsque l'une des `SourceURI` touches `FileformatAuthor`, ou correspond aux valeurs spécifiées.

```
response=kendra.query(
        QueryText = query,
        IndexId = index,
        AttributeFilter = {'OrAllFilters':
            [
                {"EqualsTo": {"Key": "Fileformat","Value": {"StringValue":
 "AUTO_DETECT"}}},
                {"EqualsTo": {"Key": "Author","Value": {"StringValue": "Ana
 Carolina"}}},
                {"EqualsTo": {"Key": "SourceURI","Value": {"StringValue": "https://
aws.amazonaws.com/234234242342"}}}
            ]
            }
        )
```

Pour `StringList` les champs, utilisez les filtres `ContainsAll` d'attributs `ContainsAny` ou pour renvoyer les documents contenant la chaîne spécifiée. L'exemple suivant montre comment renvoyer tous les documents dont l'attribut `Locations` personnalisé contient les valeurs « Seattle » ou « Portland ».

```
response=kendra.query(
        QueryText = query,
        IndexId = index,
        AttributeFilter = {
                "ContainsAny": { "Key": "Locations", "Value": { "StringListValue":
 [ "Seattle", "Portland"] }}
            }
        )
```

Filtrer les attributs de chaque document dans les résultats de recherche

Amazon Kendra renvoie les attributs de document pour chaque document figurant dans les résultats de recherche. Vous pouvez filtrer certains attributs de document que vous souhaitez inclure dans la réponse dans les résultats de recherche. Par défaut, tous les attributs de document assignés à un document sont renvoyés dans la réponse.

Dans l'exemple suivant, seuls les attributs `_source_uri` et du `_author` document sont inclus dans la réponse d'un document.

```
response=kendra.query(
        QueryText = query,
        IndexId = index,
        RequestedDocumentAttributes = ["_source_uri", "_author"]
        )
```

Filtrage en fonction du contexte utilisateur

> **Note**
>
> La prise en charge des fonctionnalités varie en fonction du type d'index et de l'API de recherche utilisés. Pour savoir si cette fonctionnalité est prise en charge pour le type d'index et l'API de recherche que vous utilisez, consultez la section Types d'index.

Vous pouvez filtrer les résultats de recherche d'un utilisateur en fonction de l'accès de celui-ci ou de son groupe aux documents. Vous pouvez utiliser un jeton utilisateur, un ID utilisateur ou un attribut utilisateur pour filtrer les documents.

Le filtrage du contexte utilisateur est une sorte de recherche personnalisée qui a l'avantage de contrôler l'accès aux documents. Par exemple, les équipes qui recherchent des informations sur le portail de l'entreprise ne doivent pas toutes accéder aux documents top secrets de l'entreprise, et ces documents ne sont pas pertinents pour tous les utilisateurs. Seuls des utilisateurs ou des groupes d'équipes spécifiques ayant accès à des documents top secrets devraient voir ces documents dans leurs résultats de recherche.

Lorsqu'un document est indexé dans Amazon Kendra, une liste de contrôle d'accès (ACL) correspondante est ingérée pour la plupart des documents. L'ACL indique les noms d'utilisateur et

de groupe auxquels l'accès au document est autorisé ou refusé. Les documents sans ACL sont des documents publics.

Amazon Kendra peut extraire les informations d'utilisateur ou de groupe associées à chaque document pour la plupart des sources de données. Par exemple, un document dans Quip peut inclure une liste « partagée » de certains utilisateurs autorisés à accéder au document. Si vous utilisez un compartiment S3 comme source de données, vous fournissez un fichier JSON pour votre ACL et vous incluez le chemin S3 vers ce fichier dans le cadre de la configuration de la source de données. Si vous ajoutez des documents directement à un index, vous spécifiez l'ACL dans l'objet Principal en tant que partie intégrante de l'objet document dans l'BatchPutDocumentAPI.

Si vous utilisez un index Amazon Kendra Enterprise ou Developer Edition, vous pouvez utiliser l'CreateAccessControlConfigurationAPI pour reconfigurer le contrôle d'accès au niveau des documents existant sans devoir indexer à nouveau tous vos documents. Par exemple, votre index contient des documents d'entreprise top secrets auxquels seuls certains employés ou utilisateurs peuvent accéder. L'un de ces utilisateurs quitte l'entreprise ou rejoint une équipe qui devrait être empêchée d'accéder à des documents top secrets. L'utilisateur a toujours accès aux documents top secrets car il y avait accès lors de l'indexation précédente de vos documents. Vous pouvez créer une configuration de contrôle d'accès spécifique pour l'utilisateur en lui refusant l'accès. Vous pouvez ultérieurement mettre à jour la configuration du contrôle d'accès pour autoriser l'accès au cas où l'utilisateur reviendrait dans l'entreprise et rejoindrait l'équipe « top-secrète ». Vous pouvez reconfigurer le contrôle d'accès à vos documents en fonction de l'évolution des circonstances.

Pour appliquer votre configuration de contrôle d'accès à certains documents, vous appelez l'BatchPutDocumentAPI avec l'objet `AccessControlConfigurationId` inclus dans le document. Si vous utilisez un compartiment S3 comme source de données, vous le mettez à jour `.metadata.json` avec la source de données `AccessControlConfigurationId` et vous le synchronisez. Amazon Kendra ne prend actuellement en charge que la configuration du contrôle d'accès pour les sources de données S3 et les documents indexés à l'aide de l'`BatchPutDocument`API.

Filtrage par jeton utilisateur

> **Important**
>
> Les indices Amazon Kendra GenAI Enterprise Edition ne prennent pas en charge le contrôle d'accès utilisateur basé sur des jetons.

Lorsque vous interrogez un index, vous pouvez utiliser un jeton utilisateur pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Lorsque vous émettez une requête, Amazon Kendra extrayez et validez le jeton, extrayez et vérifiez les informations relatives à l'utilisateur et au groupe, puis exécutez la requête. Tous les documents auxquels l'utilisateur a accès, y compris les documents publics, sont renvoyés. Pour plus d'informations, consultez la section Contrôle d'accès utilisateur basé sur des jetons.

Vous fournissez le jeton utilisateur dans l'UserContextobjet et vous le transmettez à l'API Query.

Ce qui suit montre comment inclure un jeton d'utilisateur.

```
response = kendra.query(
    QueryText = query,
    IndexId = index,
    UserToken = {
        Token = "token"
    })
```

Vous pouvez associer des utilisateurs à des groupes. Lorsque vous utilisez le filtrage contextuel utilisateur, il n'est pas nécessaire d'inclure tous les groupes auxquels appartient un utilisateur lorsque vous émettez la requête. Avec l'PutPrincipalMappingAPI, vous pouvez associer les utilisateurs à leurs groupes. Si vous ne souhaitez pas utiliser l'`PutPrincipalMapping`API, vous devez fournir le nom d'utilisateur et tous les groupes auxquels il appartient lorsque vous émettez une requête. Vous pouvez également récupérer les niveaux d'accès des groupes et des utilisateurs dans votre source d'identité IAM Identity Center à l'aide de l'UserGroupResolutionConfigurationobjet.

Filtrage par nom d'utilisateur et par groupe

Lorsque vous interrogez un index, vous pouvez utiliser l'ID utilisateur et le groupe pour filtrer les résultats de recherche en fonction de l'accès de l'utilisateur ou de son groupe aux documents. Lorsque vous émettez une requête, Amazon Kendra vérifie les informations relatives à l'utilisateur et au groupe et exécute la requête. Tous les documents relatifs à la requête auxquels l'utilisateur a accès, y compris les documents publics, sont renvoyés.

Vous pouvez également filtrer les résultats de recherche en fonction des sources de données auxquelles les utilisateurs et les groupes ont accès. La spécification d'une source de données est utile si un groupe est lié à plusieurs sources de données, mais que vous souhaitez uniquement que le groupe accède aux documents d'une certaine source de données. Par exemple, les groupes « Recherche », « Ingénierie » et « Ventes et marketing » sont tous liés aux documents de l'entreprise

stockés dans les sources de données Confluence et Salesforce. Toutefois, l'équipe « Ventes et marketing » n'a besoin d'accéder qu'aux documents relatifs aux clients stockés dans Salesforce. Ainsi, lorsque les utilisateurs des ventes et du marketing recherchent des documents relatifs aux clients, ils peuvent voir les documents de Salesforce dans leurs résultats. Les utilisateurs qui ne travaillent pas dans le domaine des ventes et du marketing ne voient pas les documents Salesforce dans leurs résultats de recherche.

Vous fournissez les informations relatives aux utilisateurs, aux groupes et aux sources de données dans l'UserContextobjet et vous les transmettez à l'API Query. L'ID utilisateur et la liste des groupes et des sources de données doivent correspondre au nom que vous spécifiez dans l'objet Principal pour identifier l'utilisateur, les groupes et les sources de données. Avec `Principal` cet objet, vous pouvez ajouter un utilisateur, un groupe ou une source de données à une liste d'autorisation ou de refus d'accès à un document.

Vous devez fournir l'un des éléments suivants :

- Informations sur les utilisateurs et les groupes, et informations (facultatives) sur les sources de données.
- Uniquement les informations utilisateur si vous mappez vos utilisateurs à des groupes et à des sources de données à l'aide de l'PutPrincipalMappingAPI. Vous pouvez également récupérer les niveaux d'accès des groupes et des utilisateurs dans votre source d'identité IAM Identity Center à l'aide de l'UserGroupResolutionConfigurationobjet.

Si ces informations ne sont pas incluses dans la requête, Amazon Kendra renvoie tous les documents. Si vous fournissez ces informations, seuls les documents avec des utilisateurs IDs, des groupes et des sources de données correspondants sont renvoyés.

Ce qui suit montre comment inclure l'ID utilisateur, les groupes et les sources de données.

```
response = kendra.query(
    QueryText = query,
    IndexId = index,
    UserContext={'Token': 'string', 'UserId': 'string', 'Groups': [ 'string', ],
 'DataSourceGroups': [ { 'GroupId': 'string', 'DataSourceId': 'string' }, ] },)
```

Filtrage par attribut utilisateur

Lorsque vous interrogez un index, vous pouvez utiliser des attributs intégrés `_user_id` et filtrer `_group_id` les résultats de recherche en fonction de l'accès de l'utilisateur et de son groupe aux

documents. Vous pouvez configurer jusqu'à 100 identifiants de groupe. Lorsque vous émettez une requête, Amazon Kendra vérifie les informations relatives à l'utilisateur et au groupe et exécute la requête. Tous les documents relatifs à la requête auxquels l'utilisateur a accès, y compris les documents publics, sont renvoyés.

Vous fournissez les attributs d'utilisateur et de groupe dans l'AttributeFilterobjet et vous les transmettez à l'API Query.

L'exemple suivant montre une demande qui filtre la réponse à la requête en fonction de l'ID utilisateur et des groupes « HR » et « IT » auxquels appartient l'utilisateur. La requête renverra tout document dont l'utilisateur ou les groupes « RH » ou « IT » figurent dans la liste des documents autorisés. Si l'utilisateur ou l'un des groupes figure dans la liste de refus d'un document, le document n'est pas renvoyé.

```
response = kendra.query(
        QueryText = query,
        IndexId = index,
        AttributeFilter = {
            "OrAllFilters": [
                {
                    "EqualsTo": {
                        "Key": "_user_id",
                        "Value": {
                            "StringValue": "user1"
                        }
                     }
                },
                {
                    "EqualsTo": {
                        "Key": "_group_ids",
                        "Value": {
                            "StringListValue": ["HR", "IT"]
                        }
                    }
                }
            ]
        }
        )
```

Vous pouvez également spécifier à quelle source de données un groupe peut accéder dans l'`Principal`objet.

> **Note**
>
> Le filtrage du contexte utilisateur n'est pas un contrôle d'authentification ou d'autorisation pour votre contenu. Il n'authentifie pas l'utilisateur et les groupes envoyés à l'`Query`API. Il appartient à votre application de s'assurer que les informations relatives aux utilisateurs et aux groupes envoyées à `Query` l'API sont authentifiées et autorisées.

Il existe une implémentation du filtrage du contexte utilisateur pour chaque source de données. La section suivante décrit chaque implémentation.

Rubriques

- Filtrage du contexte utilisateur pour les documents ajoutés directement à un index
- Filtrage du contexte utilisateur pour les questions fréquemment posées
- Filtrage du contexte utilisateur pour les sources de données

Filtrage du contexte utilisateur pour les documents ajoutés directement à un index

Lorsque vous ajoutez des documents directement à un index à l'aide de l'BatchPutDocumentAPI, les informations relatives aux utilisateurs et aux groupes sont Amazon Kendra extraites du `AccessControlList` champ du document. Vous fournissez une liste de contrôle d'accès (ACL) pour vos documents et l'ACL est ingérée avec vos documents.

Vous spécifiez l'ACL dans l'objet Principal dans le cadre de l'objet Document de l'`BatchPutDocument`API. Vous fournissez les informations suivantes :

- L'accès que l'utilisateur ou le groupe doit avoir. Tu peux dire `ALLOW` ou`DENY`.
- Type d'entité. Tu peux dire `USER` ou`GROUP`.
- Nom de l'utilisateur ou du groupe.

Vous pouvez ajouter jusqu'à 200 entrées dans le `AccessControlList` champ.

Filtrage du contexte utilisateur pour les questions fréquemment posées

Lorsque vous ajoutez une FAQ à un index, vous obtenez des informations sur Amazon Kendra les utilisateurs et les groupes à partir de l'`AccessControlList`objet/champ du fichier JSON de la FAQ. Vous pouvez également utiliser un fichier CSV de FAQ avec des champs ou des attributs personnalisés pour le contrôle d'accès.

Vous fournissez les informations suivantes :

- L'accès que l'utilisateur ou le groupe doit avoir. Tu peux dire `ALLOW` ou`DENY`.
- Type d'entité. Tu peux dire `USER` ou`GROUP`.
- Nom de l'utilisateur ou du groupe.

Pour plus d'informations, consultez les fichiers de FAQ.

Filtrage du contexte utilisateur pour les sources de données

Amazon Kendra analyse également les informations des listes de contrôle d'accès (ACL) des utilisateurs et des groupes à partir des connecteurs de source de données pris en charge. Cela est utile pour le filtrage du contexte utilisateur, où les résultats de recherche sont filtrés en fonction de l'accès de l'utilisateur ou de son groupe aux documents.

> **Important**
>
> Les indices Amazon Kendra GenAI Enterprise Edition ne prennent en charge que les connecteurs de source de données Amazon Kendra v2.0.

Rubriques

- Filtrage du contexte utilisateur pour les sources de données Adobe Experience Manager
- Filtrage du contexte utilisateur pour les sources de données Alfresco
- Filtrage du contexte utilisateur pour les sources de données Aurora (MySQL)
- Filtrage du contexte utilisateur pour les Aurora sources de données (PostgreSQL)
- Filtrage du contexte utilisateur pour les sources Amazon FSx de données
- Filtrage du contexte utilisateur pour les sources de données de base de données

- Filtrage du contexte utilisateur pour les sources de données Amazon RDS (Microsoft SQL Server)
- Filtrage du contexte utilisateur pour les sources de données Amazon RDS (MySQL)
- Filtrage du contexte utilisateur pour les sources de données Amazon RDS (Oracle)
- Filtrage du contexte utilisateur pour les Amazon RDS sources de données (PostgreSQL)
- Filtrage du contexte utilisateur pour les sources Amazon S3 de données
- Filtrage du contexte utilisateur pour les sources de données Box
- Filtrage du contexte utilisateur pour les sources de données Confluence
- Filtrage du contexte utilisateur pour les sources de données Dropbox
- Filtrage du contexte utilisateur pour les sources de données Drupal
- Filtrage du contexte utilisateur pour les sources GitHub de données
- Filtrage du contexte utilisateur pour les sources de données Gmail
- Filtrage du contexte utilisateur pour les sources de données Google Drive
- Filtrage du contexte utilisateur pour les sources DB2 de données IBM
- Filtrage du contexte utilisateur pour les sources de données Jira
- Filtrage du contexte utilisateur pour les sources de données Microsoft Exchange
- Filtrage du contexte utilisateur pour les sources OneDrive de données Microsoft
- Filtrage du contexte utilisateur pour les sources de données Microsoft OneDrive v2.0
- Filtrage du contexte utilisateur pour les sources SharePoint de données Microsoft
- Filtrage du contexte utilisateur pour les sources de données Microsoft SQL Server
- Filtrage du contexte utilisateur pour les sources de données Microsoft Teams
- Filtrage du contexte utilisateur pour les sources de données Microsoft Yammer
- Filtrage du contexte utilisateur pour les sources de données MySQL
- Filtrage du contexte utilisateur pour les sources de données Oracle Database
- Filtrage du contexte utilisateur pour les sources de données PostgreSQL
- Filtrage du contexte utilisateur pour les sources de données Quip
- Filtrage du contexte utilisateur pour les sources de données Salesforce
- Filtrage du contexte utilisateur pour les sources ServiceNow de données
- Filtrage du contexte utilisateur pour les sources de données Slack
- Filtrage du contexte utilisateur pour les sources de données Zendesk

Filtrage du contexte utilisateur pour les sources de données Adobe Experience Manager

Lorsque vous utilisez une source de données Adobe Experience Manager, vous obtenez les Amazon Kendra informations relatives aux utilisateurs et aux groupes à partir de l'instance d'Adobe Experience Manager.

Le groupe et l'utilisateur IDs sont mappés comme suit :

- `_group_ids`—Les groupes IDs existent dans le contenu d'Adobe Experience Manager pour lesquels des autorisations d'accès sont définies. Ils sont mappés à partir des noms des groupes dans Adobe Experience Manager.
- `_user_id`—L'utilisateur IDs existe dans le contenu d'Adobe Experience Manager pour lequel des autorisations d'accès sont définies. Ils sont mappés à partir des e-mails des utilisateurs, comme IDs dans Adobe Experience Manager.

Vous pouvez ajouter jusqu'à 200 entrées dans le `AccessControlList` champ.

Filtrage du contexte utilisateur pour les sources de données Alfresco

Lorsque vous utilisez une source de données Alfresco, Amazon Kendra elle obtient les informations relatives aux utilisateurs et aux groupes à partir de l'instance Alfresco.

Le groupe et l'utilisateur IDs sont mappés comme suit :

- `_group_ids`—Les groupes IDs existent dans Alfresco sur les fichiers pour lesquels des autorisations d'accès sont définies. Ils sont mappés à partir des noms système des groupes (et non des noms d'affichage) dans Alfresco.
- `_user_id`—Les utilisateurs IDs existent dans Alfresco sur les fichiers pour lesquels des autorisations d'accès sont définies. Ils sont mappés à partir des e-mails des utilisateurs comme IDs dans Alfresco.

Vous pouvez ajouter jusqu'à 200 entrées dans le `AccessControlList` champ.

Filtrage du contexte utilisateur pour les sources de données Aurora (MySQL)

Lorsque vous utilisez une source de données Aurora (MySQL), les informations relatives aux utilisateurs et aux groupes sont Amazon Kendra extraites d'une colonne de la table source. Vous

spécifiez cette colonne dans la console ou en utilisant l'TemplateConfigurationobjet dans le cadre de l'CreateDataSourceAPI.

Une source de données de base de données Aurora (MySQL) présente les limites suivantes :

- Vous pouvez uniquement spécifier une liste d'autorisations pour une source de données de base de données. Vous ne pouvez pas spécifier de liste de refus.
- Vous ne pouvez spécifier que des groupes. Vous ne pouvez pas spécifier d'utilisateurs individuels pour la liste d'autorisation.
- La colonne de base de données doit être une chaîne contenant une liste de groupes délimitée par des points-virgules.

Filtrage du contexte utilisateur pour les Aurora sources de données (PostgreSQL)

Lorsque vous utilisez une source de données Aurora (PostgreSQL) Amazon Kendra , les informations relatives aux utilisateurs et aux groupes sont extraites d'une colonne de la table source. Vous spécifiez cette colonne dans la console ou en utilisant l'TemplateConfigurationobjet dans le cadre de l'CreateDataSourceAPI.

Une source de données de base de données Aurora (PostgreSQL) présente les limites suivantes :

- Vous pouvez uniquement spécifier une liste d'autorisations pour une source de données de base de données. Vous ne pouvez pas spécifier de liste de refus.
- Vous ne pouvez spécifier que des groupes. Vous ne pouvez pas spécifier d'utilisateurs individuels pour la liste d'autorisation.
- La colonne de base de données doit être une chaîne contenant une liste de groupes délimitée par des points-virgules.

Filtrage du contexte utilisateur pour les sources Amazon FSx de données

Lorsque vous utilisez une source de Amazon FSx données, Amazon Kendra elle obtient des informations sur les utilisateurs et les groupes à partir du service d'annuaire de l' Amazon FSx instance.

Le Amazon FSx groupe et l'utilisateur IDs sont mappés comme suit :

- `_group_ids`—Le groupe IDs existe dans Amazon FSx les fichiers pour lesquels des autorisations d'accès sont définies. Ils sont mappés à partir des noms de groupes de systèmes dans le service d'annuaire de Amazon FSx.
- `_user_id`—L'utilisateur IDs existe dans Amazon FSx les fichiers pour lesquels des autorisations d'accès sont définies. Ils sont mappés à partir des noms d'utilisateur du système dans le service d'annuaire de Amazon FSx.

Vous pouvez ajouter jusqu'à 200 entrées dans le `AccessControlList` champ.

Filtrage du contexte utilisateur pour les sources de données de base de données

Lorsque vous utilisez une source de données de base de données, telle que Amazon Aurora PostgreSQL, Amazon Kendra obtient des informations sur les utilisateurs et les groupes à partir d'une colonne de la table source. Vous spécifiez cette colonne dans l'AclConfigurationobjet en tant que partie intégrante de l'DatabaseConfigurationobjet dans l'CreateDataSourceAPI.

Une source de données de base de données présente les limites suivantes :

- Vous pouvez uniquement spécifier une liste d'autorisations pour une source de données de base de données. Vous ne pouvez pas spécifier de liste de refus.
- Vous ne pouvez spécifier que des groupes. Vous ne pouvez pas spécifier d'utilisateurs individuels pour la liste d'autorisation.
- La colonne de base de données doit être une chaîne contenant une liste de groupes délimitée par des points-virgules.

Filtrage du contexte utilisateur pour les sources de données Amazon RDS (Microsoft SQL Server)

Lorsque vous utilisez une source de données Amazon RDS (Microsoft SQL Server), les informations relatives aux utilisateurs et aux groupes sont Amazon Kendra extraites d'une colonne de la table source. Vous spécifiez cette colonne dans la console ou en utilisant l'TemplateConfigurationobjet dans le cadre de l'CreateDataSourceAPI.

Une source de données de base de données Amazon RDS (Microsoft SQL Server) présente les limites suivantes :

- Vous pouvez uniquement spécifier une liste d'autorisations pour une source de données de base de données. Vous ne pouvez pas spécifier de liste de refus.

- Vous ne pouvez spécifier que des groupes. Vous ne pouvez pas spécifier d'utilisateurs individuels pour la liste d'autorisation.
- La colonne de base de données doit être une chaîne contenant une liste de groupes délimitée par des points-virgules.

Filtrage du contexte utilisateur pour les sources de données Amazon RDS (MySQL)

Lorsque vous utilisez une source de données Amazon RDS (MySQL), les informations relatives aux utilisateurs et aux groupes sont Amazon Kendra extraites d'une colonne de la table source. Vous spécifiez cette colonne dans la console ou en utilisant l'TemplateConfigurationobjet dans le cadre de l'CreateDataSourceAPI.

Une source de données de base de données Amazon RDS (MySQL) présente les limites suivantes :

- Vous pouvez uniquement spécifier une liste d'autorisations pour une source de données de base de données. Vous ne pouvez pas spécifier de liste de refus.
- Vous ne pouvez spécifier que des groupes. Vous ne pouvez pas spécifier d'utilisateurs individuels pour la liste d'autorisation.
- La colonne de base de données doit être une chaîne contenant une liste de groupes délimitée par des points-virgules.

Filtrage du contexte utilisateur pour les sources de données Amazon RDS (Oracle)

Lorsque vous utilisez une source de données Amazon RDS (Oracle), les informations relatives aux utilisateurs et aux groupes sont Amazon Kendra extraites d'une colonne de la table source. Vous spécifiez cette colonne dans la console ou en utilisant l'TemplateConfigurationobjet dans le cadre de l'CreateDataSourceAPI.

Une source de données de base de données Amazon RDS (Oracle) présente les limites suivantes :

- Vous pouvez uniquement spécifier une liste d'autorisations pour une source de données de base de données. Vous ne pouvez pas spécifier de liste de refus.
- Vous ne pouvez spécifier que des groupes. Vous ne pouvez pas spécifier d'utilisateurs individuels pour la liste d'autorisation.
- La colonne de base de données doit être une chaîne contenant une liste de groupes délimitée par des points-virgules.

Filtrage du contexte utilisateur pour les Amazon RDS sources de données (PostgreSQL)

Lorsque vous utilisez une source de données Amazon RDS (PostgreSQL) Amazon Kendra , les informations relatives aux utilisateurs et aux groupes sont extraites d'une colonne de la table source. Vous spécifiez cette colonne dans la console ou en utilisant l'TemplateConfigurationobjet dans le cadre de l'CreateDataSourceAPI.

Une source de données de base de données Amazon RDS (PostgreSQL) présente les limites suivantes :

- Vous pouvez uniquement spécifier une liste d'autorisations pour une source de données de base de données. Vous ne pouvez pas spécifier de liste de refus.
- Vous ne pouvez spécifier que des groupes. Vous ne pouvez pas spécifier d'utilisateurs individuels pour la liste d'autorisation.
- La colonne de base de données doit être une chaîne contenant une liste de groupes délimitée par des points-virgules.

Filtrage du contexte utilisateur pour les sources Amazon S3 de données

Vous ajoutez un filtrage contextuel utilisateur à un document dans une source de Amazon S3 données à l'aide d'un fichier de métadonnées associé au document. Vous ajoutez les informations dans le `AccessControlList` champ du document JSON. Pour plus d'informations sur l'ajout de métadonnées aux documents indexés à partir d'une source de Amazon S3 données, consultez la section Métadonnées des documents S3.

Vous fournissez trois informations :

- L'accès que l'entité doit avoir. Tu peux dire `ALLOW` ou`DENY`.
- Type d'entité. Tu peux dire `USER` ou`GROUP`.
- Le nom de l'entité.

Vous pouvez ajouter jusqu'à 200 entrées dans le `AccessControlList` champ.

Filtrage du contexte utilisateur pour les sources de données Box

Lorsque vous utilisez une source de données Box, Amazon Kendra vous obtenez des informations sur les utilisateurs et les groupes à partir de l'instance Box.

Le groupe Box et l'utilisateur IDs sont mappés comme suit :

- `_group_ids`—Le groupe IDs existe dans Box sur les fichiers pour lesquels des autorisations d'accès sont définies. Ils sont mappés à partir des noms des groupes dans Box.
- `_user_id`—L'utilisateur IDs existe dans Box sur les fichiers pour lesquels des autorisations d'accès sont définies. Ils sont mappés à partir des e-mails de l'utilisateur en tant qu'utilisateur IDs dans Box.

Vous pouvez ajouter jusqu'à 200 entrées dans le `AccessControlList` champ.

Filtrage du contexte utilisateur pour les sources de données Confluence

Lorsque vous utilisez une source de données Confluence, Amazon Kendra elle obtient des informations sur les utilisateurs et les groupes à partir de l'instance Confluence.

Vous configurez l'accès des utilisateurs et des groupes aux espaces à l'aide de la page des autorisations d'espace. Pour les pages et les blogs, vous utilisez la page des restrictions. Pour plus d'informations sur les autorisations d'espace, consultez la section Vue d'ensemble des autorisations d'espace sur le site Web d'assistance de Confluence. Pour plus d'informations sur les restrictions relatives aux pages et aux blogs, consultez la section Restrictions relatives aux pages sur le site Web d'assistance de Confluence.

Le groupe Confluence et les noms d'utilisateur sont mappés comme suit :

- `_group_ids`—Les noms de groupes sont présents sur les espaces, les pages et les blogs soumis à des restrictions. Ils sont mappés à partir du nom du groupe dans Confluence. Les noms de groupes sont toujours en minuscules.

- `_user_id`—Les noms d'utilisateur sont présents sur l'espace, la page ou le blog soumis à des restrictions. Ils sont mappés en fonction du type d'instance Confluence que vous utilisez.

  Pour connecteur Confluence v1.0

  - Serveur : il `_user_id` s'agit du nom d'utilisateur. Le nom d'utilisateur est toujours en minuscules.
  - Cloud : il `_user_id` s'agit de l'ID de compte de l'utilisateur.

  Pour connecteur Confluence v2.0

  - Serveur : il `_user_id` s'agit du nom d'utilisateur. Le nom d'utilisateur est toujours en minuscules.
  - Cloud : il `_user_id` s'agit de l'identifiant e-mail de l'utilisateur.

> **Important**
>
> Pour que le filtrage du contexte utilisateur fonctionne correctement pour votre connecteur Confluence, vous devez vous assurer que la visibilité d'un utilisateur autorisé à accéder à une page Confluence est définie sur Tout le monde. Pour plus d'informations, consultez la section Configurer la visibilité de vos e-mails dans la documentation Atlassian destinée aux développeurs.

Vous pouvez ajouter jusqu'à 200 entrées dans le `AccessControlList` champ.

Filtrage du contexte utilisateur pour les sources de données Dropbox

Lorsque vous utilisez une source de données Dropbox, les informations relatives aux utilisateurs et aux groupes sont Amazon Kendra extraites de l'instance Dropbox.

Le groupe et l'utilisateur IDs sont mappés comme suit :

- `_group_ids`—Les groupes IDs existent dans Dropbox sur les fichiers pour lesquels des autorisations d'accès sont définies. Ils sont mappés à partir des noms des groupes dans Dropbox.
- `_user_id`—L'utilisateur IDs existe dans Dropbox sur les fichiers pour lesquels des autorisations d'accès sont définies. Ils sont mappés à partir des e-mails des utilisateurs, comme IDs dans Dropbox.

Vous pouvez ajouter jusqu'à 200 entrées dans le `AccessControlList` champ.

Filtrage du contexte utilisateur pour les sources de données Drupal

Lorsque vous utilisez une source de données Drupal, vous obtenez les informations de l' Amazon Kendra utilisateur et du groupe à partir de l'instance Drupal.

Le groupe et l'utilisateur IDs sont mappés comme suit :

- `_group_ids`— Les groupes IDs existent dans Drupal sur les fichiers pour lesquels des autorisations d'accès sont définies. Ils sont mappés à partir des noms des groupes dans Drupal.
- `_user_id`— L'utilisateur IDs existe dans Drupal sur les fichiers pour lesquels des autorisations d'accès sont définies. Ils sont mappés à partir des e-mails des utilisateurs comme IDs dans Drupal.

Vous pouvez ajouter jusqu'à 200 entrées dans le `AccessControlList` champ.

Filtrage du contexte utilisateur pour les sources GitHub de données

Lorsque vous utilisez une source de GitHub données, Amazon Kendra obtient les informations utilisateur de l' GitHub instance.

Les GitHub utilisateurs IDs sont mappés comme suit :

- `_user_id`—L'utilisateur IDs existe dans GitHub les fichiers pour lesquels des autorisations d'accès sont définies. Ils sont mappés à partir des e-mails des utilisateurs sous forme d'entrée IDs . GitHub

Vous pouvez ajouter jusqu'à 200 entrées dans le `AccessControlList` champ.

Filtrage du contexte utilisateur pour les sources de données Gmail

Lorsque vous utilisez une source de données Gmail, Amazon Kendra obtient les informations utilisateur de l'instance Gmail.

Les utilisateurs IDs sont mappés comme suit :

- `_user_id`— L'utilisateur IDs existe dans Gmail sur les fichiers pour lesquels des autorisations d'accès sont définies. Ils sont mappés à partir des e-mails des utilisateurs comme IDs dans Gmail.

Vous pouvez ajouter jusqu'à 200 entrées dans le `AccessControlList` champ.

Filtrage du contexte utilisateur pour les sources de données Google Drive

Une source de données Google Workspace Drive renvoie des informations sur les utilisateurs et les groupes relatifs aux utilisateurs et aux groupes de Google Drive. L'appartenance au groupe et au domaine est mappée au champ d'`_group_ids`index. Le nom d'utilisateur Google Drive est associé au `_user_id` champ.

Lorsque vous fournissez une ou plusieurs adresses e-mail utilisateur dans l'`Query`API, seuls les documents partagés avec ces adresses e-mail sont renvoyés. Le `AttributeFilter` paramètre suivant renvoie uniquement les documents partagés avec "martha@example.com ».

```
"AttributeFilter": {
```

```
                "EqualsTo":{
                   "Key": "_user_id",
                   "Value": {
                       "StringValue": "martha@example.com"
                   }
               }
           }
```

Si vous fournissez une ou plusieurs adresses e-mail de groupe dans la requête, seuls les documents partagés avec les groupes sont renvoyés. Le `AttributeFilter` paramètre suivant renvoie uniquement les documents partagés avec le groupe « hr@example.com ».

```
"AttributeFilter": {
                "EqualsTo":{
                   "Key": "_group_ids",
                   "Value": {
                       "StringListValue": ["hr@example.com"]
                   }
               }
           }
```

Si vous indiquez le domaine dans la requête, tous les documents partagés avec le domaine sont renvoyés. Le `AttributeFilter` paramètre suivant renvoie les documents partagés avec le domaine « exemple.com ».

```
"AttributeFilter": {
                "EqualsTo":{
                   "Key": "_group_ids",
                   "Value": {
                       "StringListValue": ["example.com"]
                   }
               }
           }
```

Vous pouvez ajouter jusqu'à 200 entrées dans le `AccessControlList` champ.

Filtrage du contexte utilisateur pour les sources DB2 de données IBM

Lorsque vous utilisez une source de DB2 données IBM, vous obtenez Amazon Kendra des informations sur les utilisateurs et les groupes à partir d'une colonne de la table source. Vous

spécifiez cette colonne dans la console ou en utilisant l'TemplateConfigurationobjet dans le cadre de l'CreateDataSourceAPI.

Une source DB2 de données de base de données IBM présente les limites suivantes :

- Vous pouvez uniquement spécifier une liste d'autorisations pour une source de données de base de données. Vous ne pouvez pas spécifier de liste de refus.
- Vous ne pouvez spécifier que des groupes. Vous ne pouvez pas spécifier d'utilisateurs individuels pour la liste d'autorisation.
- La colonne de base de données doit être une chaîne contenant une liste de groupes délimitée par des points-virgules.

Filtrage du contexte utilisateur pour les sources de données Jira

Lorsque vous utilisez une source de données Jira, Amazon Kendra elle obtient des informations sur les utilisateurs et les groupes à partir de l'instance Jira.

Les utilisateurs Jira IDs sont mappés comme suit :

- `_user_id`—L'utilisateur IDs existe dans Jira sur les fichiers pour lesquels des autorisations d'accès sont définies. Ils sont mappés à partir des e-mails de l'utilisateur en tant qu'utilisateur IDs dans Jira.

Vous pouvez ajouter jusqu'à 200 entrées dans le `AccessControlList` champ.

Filtrage du contexte utilisateur pour les sources de données Microsoft Exchange

Lorsque vous utilisez une source de données Microsoft Exchange, Amazon Kendra obtient les informations utilisateur à partir de l'instance Microsoft Exchange.

Les utilisateurs Microsoft Exchange IDs sont mappés comme suit :

- `_user_id`—L'utilisateur IDs existe dans Microsoft Exchange des autorisations lui permettant d'accéder à certains contenus. Ils sont mappés à partir des noms d'utilisateur comme IDs dans Microsoft Exchange.

Vous pouvez ajouter jusqu'à 200 entrées dans le `AccessControlList` champ.

Filtrage du contexte utilisateur pour les sources OneDrive de données Microsoft

Amazon Kendra récupère les informations relatives aux utilisateurs et aux groupes auprès de Microsoft OneDrive lorsqu'il indexe les documents du site. Les informations relatives aux utilisateurs et aux groupes sont extraites du SharePoint site Microsoft sous-jacent qui les héberge OneDrive.

Lorsque vous utilisez un OneDrive utilisateur ou un groupe pour filtrer les résultats de recherche, calculez l'ID comme suit :

1. Obtenez le nom du site. Par exemple, `https://host.onmicrosoft.com/sites/siteName`.
2. Prenez le MD5 hachage du nom du site. Par exemple, `430a6b90503eef95c89295c8999c7981`.
3. Créez l'adresse e-mail ou l'identifiant de groupe de l'utilisateur en concaténant le MD5 hachage avec une barre verticale (|) et l'identifiant. Par exemple, si le nom d'un groupe est localGroupName « », l'ID du groupe serait :

   `"430a6b90503eef95c89295c8999c7981 | localGroupName"`

   > **Note**
   >
   > Incluez un espace avant et après la barre verticale. La barre verticale est utilisée pour s'identifier `localGroupName` grâce à son MD5 hachage.

   Pour le nom d'utilisateur « someone@host.onmicrosoft.com », l'ID utilisateur serait le suivant :

   `"430a6b90503eef95c89295c8999c7981 | someone@host.onmicrosoft.com"`

Envoyez l'ID d'utilisateur ou de groupe Amazon Kendra sous forme d'`_group_id`attribut `_user_id` or lorsque vous appelez l'API Query. Par exemple, la AWS CLI commande qui utilise un groupe pour filtrer les résultats de recherche ressemble à ceci :

```
aws kendra  query \
                --index-id index ID
                --query-text "query text"
                --attribute-filter '{
                   "EqualsTo":{
                     "Key": "_group_id",
                     "Value": {"StringValue": "430a6b90503eef95c89295c8999c7981 |
 localGroupName"}
```

```
                }}'
```

Vous pouvez ajouter jusqu'à 200 entrées dans le `AccessControlList` champ.

Filtrage du contexte utilisateur pour les sources de données Microsoft OneDrive v2.0

Une source de données Microsoft OneDrive v2.0 renvoie des informations de section et de page à partir d'entités de liste de contrôle d' OneDrive accès (ACL). Amazon Kendra utilise le domaine du OneDrive locataire pour se connecter à l' OneDrive instance, puis peut filtrer les résultats de recherche en fonction de l'accès des utilisateurs ou des groupes aux sections et aux noms de fichiers.

Pour les objets standard, les `_user_id` et `_group_id` sont utilisés comme suit :

- `_user_id`— Votre adresse e-mail OneDrive utilisateur Microsoft est mappée `_user_id` sur le champ.
- `_group_id`— L'e-mail de votre OneDrive groupe Microsoft est mappé `_group_id` sur le champ.

Vous pouvez ajouter jusqu'à 200 entrées dans le `AccessControlList` champ.

Filtrage du contexte utilisateur pour les sources SharePoint de données Microsoft

Amazon Kendra récupère les informations relatives aux utilisateurs et aux groupes auprès de Microsoft SharePoint lorsqu'il indexe les documents du site. Pour filtrer les résultats de recherche en fonction de l'accès des utilisateurs ou des groupes, fournissez des informations sur les utilisateurs et les groupes lorsque vous appelez l'`Query`API.

Pour filtrer à l'aide d'un nom d'utilisateur, utilisez son adresse e-mail. Par exemple, johnstiles@example.com.

Lorsque vous utilisez un SharePoint groupe pour filtrer les résultats de recherche, calculez l'ID du groupe comme suit :

Pour les groupes locaux

1. Obtenez le nom du site. Par exemple, `https://host.onmicrosoft.com/sites/siteName`.
2. Prenez le SHA256 hachage du nom du site. Par exemple, `430a6b90503eef95c89295c8999c7981`.

3. Créez l'ID du groupe en concaténant le SHA256 hachage avec une barre verticale (|) et le nom du groupe. Par exemple, si le nom du groupe est localGroupName « », l'ID du groupe serait :

   `"430a6b90503eef95c89295c8999c7981 | localGroupName"`

   > **Note**
   >
   > Incluez un espace avant et après la barre verticale. La barre verticale est utilisée pour s'identifier `localGroupName` grâce à son SHA256 hachage.

Envoyez l'ID de groupe en Amazon Kendra tant qu'`_group_id`attribut lorsque vous appelez l'API Query. Par exemple, la AWS CLI commande ressemble à ceci :

```
aws kendra  query \
                --index-id index ID
                --query-text "query text"
                --attribute-filter '{
                   "EqualsTo":{
                     "Key": "_group_id",
                     "Value": {"StringValue": "430a6b90503eef95c89295c8999c7981 |
 localGroupName"}
                  }}'
```

Pour les groupes AD

1. Utilisez l'ID du groupe AD pour configurer le filtrage des résultats de recherche.

Envoyez l'ID de groupe en Amazon Kendra tant qu'`_group_id`attribut lorsque vous appelez l'API Query. Par exemple, la AWS CLI commande ressemble à ceci :

```
aws kendra  query \
                --index-id index ID
                --query-text "query text"
                --attribute-filter '{
                   "EqualsTo":{
                     "Key": "_group_id",
                     "Value": {"StringValue": "AD group"}
                  }}'
```

Vous pouvez ajouter jusqu'à 200 entrées dans le `AccessControlList` champ.

Filtrage du contexte utilisateur pour les sources de données Microsoft SQL Server

Lorsque vous utilisez une source de données Microsoft SQL Server, Amazon Kendra vous obtenez des informations sur les utilisateurs et les groupes à partir d'une colonne de la table source. Vous spécifiez cette colonne dans la console ou en utilisant l'TemplateConfigurationobjet dans le cadre de l'CreateDataSourceAPI.

Une source de données de base de données Microsoft SQL Server présente les limites suivantes :

- Vous pouvez uniquement spécifier une liste d'autorisations pour une source de données de base de données. Vous ne pouvez pas spécifier de liste de refus.
- Vous ne pouvez spécifier que des groupes. Vous ne pouvez pas spécifier d'utilisateurs individuels pour la liste d'autorisation.
- La colonne de base de données doit être une chaîne contenant une liste de groupes délimitée par des points-virgules.

Filtrage du contexte utilisateur pour les sources de données Microsoft Teams

Amazon Kendra récupère les informations utilisateur de Microsoft Teams lorsqu'il indexe les documents. Les informations utilisateur sont extraites de l'instance Microsoft Teams sous-jacente.

Vous pouvez ajouter jusqu'à 200 entrées dans le `AccessControlList` champ.

Filtrage du contexte utilisateur pour les sources de données Microsoft Yammer

Amazon Kendra récupère les informations utilisateur de Microsoft Yammer lorsqu'il indexe les documents. Les informations relatives à l'utilisateur et au groupe sont extraites de l'instance Microsoft Yammer sous-jacente.

Les utilisateurs de Microsoft Yammer IDs sont mappés comme suit :

- `_email_id`— L'adresse e-mail Microsoft mappée au `_user_id` champ.

Vous pouvez ajouter jusqu'à 200 entrées dans le `AccessControlList` champ.

Filtrage du contexte utilisateur pour les sources de données MySQL

Lorsque vous utilisez une source de données MySQL, les informations relatives aux utilisateurs et aux groupes sont Amazon Kendra extraites d'une colonne de la table source. Vous spécifiez cette colonne dans la console ou en utilisant l'TemplateConfigurationobjet dans le cadre de l'CreateDataSourceAPI.

Une source de données de base de données MySQL présente les limites suivantes :

- Vous pouvez uniquement spécifier une liste d'autorisations pour une source de données de base de données. Vous ne pouvez pas spécifier de liste de refus.
- Vous ne pouvez spécifier que des groupes. Vous ne pouvez pas spécifier d'utilisateurs individuels pour la liste d'autorisation.
- La colonne de base de données doit être une chaîne contenant une liste de groupes délimitée par des points-virgules.

Filtrage du contexte utilisateur pour les sources de données Oracle Database

Lorsque vous utilisez une source de données Oracle Database, les informations relatives aux utilisateurs et aux groupes sont Amazon Kendra extraites d'une colonne de la table source. Vous spécifiez cette colonne dans la console ou en utilisant l'TemplateConfigurationobjet dans le cadre de l'CreateDataSourceAPI.

Une source de données de base de données Oracle Database présente les limites suivantes :

- Vous pouvez uniquement spécifier une liste d'autorisations pour une source de données de base de données. Vous ne pouvez pas spécifier de liste de refus.
- Vous ne pouvez spécifier que des groupes. Vous ne pouvez pas spécifier d'utilisateurs individuels pour la liste d'autorisation.
- La colonne de base de données doit être une chaîne contenant une liste de groupes délimitée par des points-virgules.

Filtrage du contexte utilisateur pour les sources de données PostgreSQL

Lorsque vous utilisez une source de données PostgreSQL Amazon Kendra , les informations relatives aux utilisateurs et aux groupes sont extraites d'une colonne de la table source. Vous spécifiez cette colonne dans la console ou en utilisant l'TemplateConfigurationobjet dans le cadre de l'CreateDataSourceAPI.

Une source de données de base de données PostgreSQL présente les limites suivantes :

- Vous pouvez uniquement spécifier une liste d'autorisations pour une source de données de base de données. Vous ne pouvez pas spécifier de liste de refus.
- Vous ne pouvez spécifier que des groupes. Vous ne pouvez pas spécifier d'utilisateurs individuels pour la liste d'autorisation.
- La colonne de base de données doit être une chaîne contenant une liste de groupes délimitée par des points-virgules.

Filtrage du contexte utilisateur pour les sources de données Quip

Lorsque vous utilisez une source de données Quip, Amazon Kendra elle obtient les informations utilisateur de l'instance Quip.

Les utilisateurs de Quip IDs sont mappés comme suit :

- `_user_id`—L'utilisateur IDs existe dans Quip sur les fichiers pour lesquels des autorisations d'accès sont définies. Ils sont mappés à partir des e-mails des utilisateurs, comme IDs dans Quip.

Vous pouvez ajouter jusqu'à 200 entrées dans le `AccessControlList` champ.

Filtrage du contexte utilisateur pour les sources de données Salesforce

Une source de données Salesforce renvoie des informations sur les utilisateurs et les groupes à partir des entités de la liste de contrôle d'accès (ACL) Salesforce. Vous pouvez appliquer le filtrage du contexte utilisateur aux objets standard et aux flux de discussion Salesforce. Le filtrage du contexte utilisateur n'est pas disponible pour les articles de connaissances de Salesforce.

Si vous associez un champ Salesforce aux champs du titre et du corps du document Amazon Kendra, Amazon Kendra utilisera les données du titre et des champs du corps du document dans les réponses de recherche.

Pour les objets standard, les `_user_id` et `_group_ids` sont utilisés comme suit :

- `_user_id`: le nom d'utilisateur de l'utilisateur Salesforce.
- `_group_ids`—
  - Nom de la Salesforce `Profile`
  - Nom de la Salesforce `Group`

- Nom de la Salesforce `UserRole`
- Nom de la Salesforce `PermissionSet`

Pour les fils de discussion, les `_user_id` et `_group_ids` sont utilisés comme suit :

- `_user_id`: le nom d'utilisateur de l'utilisateur Salesforce. Disponible uniquement si l'article est publié dans le fil de l'utilisateur.
- `_group_ids` IDs —Les groupes sont utilisés comme suit. Disponible uniquement si l'élément du fil est publié dans un groupe de discussion ou de collaboration.
  - Le nom du chatteur ou du groupe de collaboration.
  - Si le groupe est public,`PUBLIC:ALL`.

Vous pouvez ajouter jusqu'à 200 entrées dans le `AccessControlList` champ.

Filtrage du contexte utilisateur pour les sources ServiceNow de données

Le filtrage du contexte utilisateur pour n' ServiceNow est pris en charge que pour l' TemplateConfiguration API et le ServiceNow Connector v2.0. ServiceNowConfiguration L'API et le ServiceNow connecteur v1.0 ne prennent pas en charge le filtrage du contexte utilisateur.

Lorsque vous utilisez une source de ServiceNow données, Amazon Kendra elle obtient les informations relatives aux utilisateurs et aux groupes à partir de l' ServiceNow instance.

Le groupe et l'utilisateur IDs sont mappés comme suit :

- `_group_ids`—Le groupe IDs existe dans ServiceNow les fichiers pour lesquels des autorisations d'accès sont définies. Ils sont mappés à partir des noms de rôles de `sys_ids` in ServiceNow.
- `_user_id`—L'utilisateur IDs existe dans ServiceNow les fichiers pour lesquels des autorisations d'accès sont définies. Ils sont mappés à partir des e-mails des utilisateurs sous forme d'entrée IDs . ServiceNow

Vous pouvez ajouter jusqu'à 200 entrées dans le `AccessControlList` champ.

Filtrage du contexte utilisateur pour les sources de données Slack

Lorsque vous utilisez une source de données Slack, Amazon Kendra elle obtient les informations utilisateur de l'instance Slack.

Les utilisateurs de Slack IDs sont mappés comme suit :

- `_user_id`—L'utilisateur IDs existe dans Slack sur les messages et les chaînes pour lesquels des autorisations d'accès sont définies. Ils sont mappés à partir des e-mails des utilisateurs, comme IDs dans Slack.

Vous pouvez ajouter jusqu'à 200 entrées dans le `AccessControlList` champ.

Filtrage du contexte utilisateur pour les sources de données Zendesk

Lorsque vous utilisez une source de données Zendesk, vous obtenez les informations Amazon Kendra de l'utilisateur et du groupe à partir de l'instance Zendesk.

Le groupe et l'utilisateur IDs sont mappés comme suit :

- `_group_ids`—Les groupes IDs existent dans les tickets et les articles Zendesk pour lesquels des autorisations d'accès sont définies. Ils sont mappés à partir des noms des groupes dans Zendesk.
- `_user_id`—Les groupes IDs existent dans les tickets et les articles Zendesk pour lesquels des autorisations d'accès sont définies. Ils sont mappés à partir des e-mails des utilisateurs, comme IDs dans Zendesk.

Vous pouvez ajouter jusqu'à 200 entrées dans le `AccessControlList` champ.

Réponses aux requêtes et types de réponses

> **Note**
>
> La prise en charge des fonctionnalités varie en fonction du type d'index et de l'API de recherche utilisés. Pour savoir si cette fonctionnalité est prise en charge pour le type d'index et l'API de recherche que vous utilisez, consultez la section Types d'index.

Amazon Kendra prend en charge différentes réponses aux requêtes et types de réponses.

Réponses aux requêtes

Un appel à l'API Query renvoie des informations sur les résultats d'une recherche. Les résultats se trouvent dans un tableau d'QueryResultItemobjets (`ResultItems`). Chacune `QueryResultItem` inclut un résumé du résultat. Les attributs de document associés au résultat de la requête sont inclus.

Informations récapitulatives

Les informations récapitulatives varient en fonction du type de résultat. Dans chaque cas, il inclut le texte du document qui correspond au terme de recherche. Il inclut également des informations de surlignage que vous pouvez utiliser pour surligner le texte de recherche dans le résultat de votre application. Par exemple, si le terme de recherche est quelle est la hauteur de la Space Needle ? , les informations récapitulatives incluent l'emplacement du texte pour les mots hauteur et aiguille spatiale. Pour plus d'informations sur les types de réponse, consultezRéponses aux requêtes et types de réponses.

Attributs du document

Chaque résultat contient les attributs du document correspondant à une requête. Certains attributs sont prédéfinis, tels que `DocumentIdDocumentTitle`, et`DocumentUri`. Les autres sont des attributs personnalisés que vous définissez. Vous pouvez utiliser les attributs du document pour filtrer la réponse provenant de l'`Query`API. Par exemple, il se peut que vous souhaitiez uniquement les documents rédigés par un auteur spécifique ou une version spécifique d'un document. Pour de plus amples informations, veuillez consulter Filtrage et recherche par facettes. Vous spécifiez les attributs du document lorsque vous ajoutez des documents à un index. Pour plus d'informations, consultez la section Champs ou attributs personnalisés.

Voici un exemple de code JSON pour le résultat d'une requête. Notez les attributs du document dans `DocumentAttributes` et`AdditionalAttributes`.

```
{
    "QueryId": "query-id",
    "ResultItems": [
        {
            "Id": "result-id",
            "Type": "ANSWER",
            "AdditionalAttributes": [
                {
                    "Key": "AnswerText",
                    "ValueType": "TEXT_WITH_HIGHLIGHTS_VALUE",
                    "Value": {
```

```
                    "TextWithHighlightsValue": {
                        "Text": "text",
                        "Highlights": [
                            {
                                "BeginOffset": 55,
                                "EndOffset": 90,
                                "TopAnswer": false
                            }
                        ]
                    }
                }
            }
        ],
        "DocumentId": "document-id",
        "DocumentTitle": {
            "Text": "title"
        },
        "DocumentExcerpt": {
            "Text": "text",
            "Highlights": [
                {
                    "BeginOffset": 0,
                    "EndOffset": 300,
                    "TopAnswer": false
                }
            ]
        },
        "DocumentURI": "uri",
        "DocumentAttributes": [],
        "ScoreAttributes": "score",
        "FeedbackToken": "token"
    },
    {
        "Id": "result-id",
        "Type": "ANSWER",
        "Format": "TABLE",
        "DocumentId": "document-id",
        "DocumentTitle": {
            "Text": "title"
        },
        "TableExcerpt": {
            "Rows": [{
                "Cells": [{
                    "Header": true,
```

```
                    "Highlighted": false,
                    "TopAnswer": false,
                    "Value": "value"
                }, {
                    "Header": true,
                    "Highlighted": false,
                    "TopAnswer": false,
                    "Value": "value"
                }, {
                    "Header": true,
                    "Highlighted": false,
                    "TopAnswer": false,
                    "Value": "value"
                }, {
                    "Header": true,
                    "Highlighted": false,
                    "TopAnswer": false,
                    "Value": "value"
                }]
            }, {
                "Cells": [{
                    "Header": false,
                    "Highlighted": false,
                    "TopAnswer": false,
                    "Value": "value"
                }, {
                    "Header": false,
                    "Highlighted": false,
                    "TopAnswer": false,
                    "Value": "value"
                }, {
                    "Header": false,
                    "Highlighted": true,
                    "TopAnswer": true,
                    "Value": "value"
                }, {
                    "Header": false,
                    "Highlighted": false,
                    "TopAnswer": false,
                    "Value": "value"
                ]}
              }],
                "TotalNumberofRows": number
    },
```

```
            "DocumentURI": "uri",
            "ScoreAttributes": "score",
            "FeedbackToken": "token"
        },
        {
            "Id": "result-id",
            "Type": "DOCUMENT",
            "AdditionalAttributes": [],
            "DocumentId": "document-id",
            "DocumentTitle": {
                "Text": "title",
                "Highlights": []
            },
            "DocumentExcerpt": {
                "Text": "text",
                "Highlights": [
                    {
                        "BeginOffset": 74,
                        "EndOffset": 77,
                        "TopAnswer": false
                    }
                ]
            },
            "DocumentURI": "uri",
            "DocumentAttributes": [
                {
                    "Key": "_source_uri",
                    "Value": {
                        "StringValue": "uri"
                    }
                }
            ],
            "ScoreAttributes": "score",
            "FeedbackToken": "token",
        }
    ],
    "FacetResults": [],
    "TotalNumberOfResults": number
}
```

Types de réponses

Amazon Kendra renvoie trois types de réponses à une requête.

- Réponse (inclut la réponse sous forme de tableau)
- Document
- Question/réponse

Le type de réponse est renvoyé dans le champ de `Type` réponse de l'QueryResultItemobjet.

Réponse

Amazon Kendra a détecté une ou plusieurs réponses à une question dans la réponse. Un factoïde est la réponse à une question qui, quoi, quand ou où, telle que Où se trouve le centre de service le plus proche de chez moi ? Amazon Kendra renvoie le texte dans l'index qui correspond le mieux à la requête. Le texte se trouve dans le `AnswerText` champ et contient des informations surlignées pour le terme de recherche dans le texte de réponse. `AnswerText`inclut l'extrait complet du document avec le texte surligné, tandis qu'`DocumentExcerpt`inclut l'extrait tronqué (290 caractères) avec le texte surligné.

Amazon Kendra ne renvoie qu'une seule réponse par document, et c'est la réponse la plus fiable. Pour renvoyer plusieurs réponses à partir d'un document, vous devez diviser le document en plusieurs documents.

```
{
    'AnswerText': {
        'TextWithHighlights': [
            {
                'BeginOffset': 271,
                'EndOffset': 279,
                'TopAnswer': False
            },
            {
                'BeginOffset': 481,
                'EndOffset': 489,
                'TopAnswer': False
            },
            {
                'BeginOffset': 547,
                'EndOffset': 555,
                'TopAnswer': False
            },
            {
                'BeginOffset': 764,
                'EndOffset': 772,
```

```
                'TopAnswer': False
            }
        ],
        'Text': 'Asynchronousoperationscan\n''alsoprocess
\n''documentsthatareinPDF''format.UsingPDFformatfilesallowsyoutoprocess''multi-
page\n''documents.\n''Forinformationabouthow''AmazonTextractrepresents
\n''documentsasBlockobjects,
        ''seeDocumentsandBlockObjects.
\n''\n''\n''\n''Forinformationaboutdocument''limits,
        seeLimitsinAmazonTextract.
\n''\n''\n''\n''TheAmazonTextractsynchronous''operationscanprocessdocumentsstoredinanAmazon
\n''S3Bucketoryoucanpass''base64encodedimagebytes.\n''Formoreinformation,

 see''CallingAmazonTextractSynchronousOperations.''Asynchronousoperationsrequireinputdocuments
\n''tobesuppliedinanAmazon''S3Bucket.'
    },
    'DocumentExcerpt': {
        'Highlights': [
            {
                'BeginOffset': 0,
                'EndOffset': 300,
                'TopAnswer': False
            }
        ],
        'Text': 'Asynchronousoperationscan\n''alsoprocess
\n''documentsthatareinPDF''format.UsingPDFformatfilesallowsyoutoprocess''multi-page
\n''documents.\n''ForinformationabouthowAmazon''Textractrepresents\n'''
    },
    'Type': 'ANSWER'
}
```

Document

Amazon Kendra renvoie les documents classés pour ceux qui correspondent au terme de recherche. Le classement est basé sur la confiance accordée Amazon Kendra à l'exactitude des résultats de recherche. Les informations relatives au document correspondant sont renvoyées dans le QueryResultItem. Il inclut le titre du document. L'extrait inclut des informations de surlignage pour le texte de recherche et la section du texte correspondant dans le document. L'URI pour les documents correspondants se trouve dans l'attribut `SourceURI` document. L'exemple de code JSON suivant montre le résumé d'un document correspondant.

```
{
```

```
    'DocumentTitle': {
        'Highlights': [
            {
                'BeginOffset': 7,
                'EndOffset': 15,
                'TopAnswer': False
            },
            {
                'BeginOffset': 97,
                'EndOffset': 105,
                'TopAnswer': False
            }
        ],
        'Text': 'AmazonTextractAPIPermissions: Actions,
        \n''Permissions,
        andResourcesReference-''AmazonTextract'
    },
    'DocumentExcerpt': {
        'Highlights': [
            {
                'BeginOffset': 68,
                'EndOffset': 76,
                'TopAnswer': False
            },
            {
                'BeginOffset': 121,
                'EndOffset': 129,
                'TopAnswer': False
            }
        ],
        'Text': '...LoggingandMonitoring\tMonitoring
\n''\tCloudWatchMetricsforAmazonTextract
\n''\tLoggingAmazonTextractAPICallswithAWSCloudTrail\n''\tAPIReference\tActions
\tAnalyzeDocument\n''\tDetectDocumentText\n''\tGetDocumentAnalysis...'
    },
    'Type': 'DOCUMENT'
}
```

Question/réponse

Une réponse aux questions et réponses est renvoyée lorsqu'une question Amazon Kendra correspond à l'une des questions fréquemment posées de votre index. La réponse inclut la question et la réponse correspondantes dans le QueryResultItemchamp. Il inclut également des informations

de surlignage pour les termes de requête détectés dans la chaîne de requête. Le JSON suivant montre une réponse à une question et à une réponse. Notez que la réponse inclut le texte de la question.

```
{
    'AnswerText': {
        'TextWithHighlights': [

        ],
        'Text': '605feet'
    },
    'DocumentExcerpt': {
        'Highlights': [
            {
                'BeginOffset': 0,
                'EndOffset': 8,
                'TopAnswer': False
            }
        ],
        'Text': '605feet'
    },
    'Type': 'QUESTION_ANSWER',
    'QuestionText': {
        'Highlights': [
            {
                'BeginOffset': 12,
                'EndOffset': 18,
                'TopAnswer': False
            },
            {
                'BeginOffset': 26,
                'EndOffset': 31,
                'TopAnswer': False
            },
            {
                'BeginOffset': 32,
                'EndOffset': 38,
                'TopAnswer': False
            }
        ],
        'Text': 'whatistheheightoftheSpaceNeedle?'
    }
}
```

Pour plus d'informations sur l'ajout de texte de question et de réponse à un index, consultez la section Création d'une FAQ.

Réglage et tri des réponses

> **Note**
>
> La prise en charge des fonctionnalités varie en fonction du type d'index et de l'API de recherche utilisés. Pour savoir si cette fonctionnalité est prise en charge pour le type d'index et l'API de recherche que vous utilisez, consultez la section Types d'index.

Vous pouvez modifier l'effet d'un champ ou d'un attribut sur la pertinence de la recherche en ajustant la pertinence. Vous pouvez également trier les résultats de recherche en fonction d'un attribut ou d'un champ spécifique.

Rubriques

- Réglage des réponses
- Tri des réponses

Réglage des réponses

Vous pouvez modifier l'effet d'un champ ou d'un attribut sur la pertinence de la recherche en ajustant la pertinence. Pour tester rapidement le réglage de la pertinence, utilisez l'API Query pour transmettre les configurations de réglage dans la requête. Ensuite, vous pouvez voir les différents résultats de recherche que vous obtenez à partir de différentes configurations. Le réglage de la pertinence au niveau de la requête n'est pas pris en charge dans la console. Vous pouvez également régler des champs ou des attributs du type défini uniquement `StringList` au niveau de l'index. Pour plus d'informations, consultez la section Optimisation de la pertinence de la recherche.

Par défaut, les réponses aux requêtes sont triées en fonction du score de pertinence qui Amazon Kendra détermine chaque résultat de la réponse.

Vous pouvez ajuster les résultats pour n'importe quel attribut/champ intégré ou personnalisé des types suivants :

- Valeur de date

- Valeur longue
- Valeur de chaîne

Vous ne pouvez pas trier les attributs du type suivant :

- Valeurs de la liste de chaînes

Classement et optimisation des résultats des documents (AWS SDK)

Définissez le `Searchable` paramètre sur true pour améliorer la configuration des métadonnées du document.

Pour régler un attribut dans une requête, définissez le `DocumentRelevanceOverrideConfigurations` paramètre de l'`Query`API et spécifiez le nom de l'attribut à régler.

L'exemple JSON suivant montre un `DocumentRelevanceOverrideConfigurations` objet qui remplace le réglage de l'attribut appelé « department » dans l'index.

```
"DocumentRelevanceOverrideConfigurations" : [
    "Name": "department",
    "Relevance": {
        "Importance": 1,
        "ValueImportanceMap": {
            "IT": 3,
            "HR": 7
        }
    }
]
```

Tri des réponses

Amazon Kendra utilise l'attribut ou le champ de tri dans le cadre des critères pour les documents renvoyés par la requête. Par exemple, les résultats renvoyés par une requête triée par « _created_at » peuvent ne pas contenir les mêmes résultats qu'une requête triée par « _version ».

Par défaut, les réponses aux requêtes sont triées en fonction du score de pertinence qui Amazon Kendra détermine chaque résultat de la réponse. Pour modifier l'ordre de tri, rendez un attribut de document triable, puis configurez Amazon Kendra pour utiliser cet attribut pour trier les réponses.

Vous pouvez trier les résultats selon n'importe quel attribut/champ intégré ou personnalisé des types suivants :

- Valeur de date
- Valeur longue
- Valeur de chaîne

Vous ne pouvez pas trier les attributs du type suivant :

- Valeurs de la liste de chaînes

Vous pouvez effectuer un tri en fonction d'un ou de plusieurs attributs de document dans chaque requête. Les requêtes renvoient 100 résultats. Si l'attribut de tri est défini dans moins de 100 documents, les documents sans valeur pour l'attribut de tri sont renvoyés à la fin des résultats, triés en fonction de leur pertinence par rapport à la requête.

Pour trier les résultats d'un document (AWS SDK)

1. Pour utiliser l'UpdateIndexAPI afin de rendre un attribut triable, définissez le `Sortable` paramètre sur. `true` L'exemple JSON suivant permet `DocumentMetadataConfigurationUpdates` d'ajouter un attribut appelé « Department » à l'index et de le rendre triable.

   ```
   "DocumentMetadataConfigurationUpdates": [
      {
          "Name": "Department",
          "Type": "STRING_VALUE",
          "Search": {
              "Sortable": "true"
          }
      }
   ]
   ```

2. Pour utiliser un attribut triable dans une requête, définissez le `SortingConfiguration` paramètre de l'API Query. Spécifiez le nom de l'attribut à trier et indiquez si vous souhaitez trier la réponse par ordre croissant ou décroissant.

   L'exemple JSON suivant montre le `SortingConfiguration` paramètre que vous utilisez pour trier les résultats d'une requête par l'attribut « Department » dans l'ordre croissant.

```
   "SortingConfiguration": {
      "DocumentAttributeKey": "Department",
      "SortOrder": "ASC"
   }
```

3. Pour utiliser plusieurs attributs triables dans une requête, définissez le `SortingConfigurations` paramètre de l'API Query. Vous pouvez définir jusqu'à 3 champs sur lesquels les résultats Amazon Kendra doivent être triés. Vous pouvez également spécifier si les résultats doivent être triés par ordre croissant ou décroissant. Le quota de champs de tri peut être augmenté.

   Si vous ne fournissez pas de configuration de tri, les résultats sont triés selon la pertinence qui Amazon Kendra détermine le résultat. En cas d'égalité dans le tri des résultats, les résultats sont triés par pertinence.

   L'exemple JSON suivant montre le `SortingConfigurations` paramètre que vous utilisez pour trier les résultats d'une requête par les attributs « Nom » et « Prix » par ordre croissant.

```
"CollapseConfiguration" : {
    "DocumentAttributeKey": "Name",
    "SortingConfigurations": [
        {
            "DocumentAttributeKey": "Price",
            "SortOrder": "ASC"
        }
    ],
    "MissingAttributeKeyStrategy": "IGNORE"
}
```

Pour trier les résultats d'un document (console)

> **Note**
>
> Le tri multi-attributs n'est actuellement pas pris en charge par le AWS Management Console.

1. Pour qu'un attribut puisse être trié dans la console, choisissez Sortable dans la définition de l'attribut. Vous pouvez rendre un attribut triable lorsque vous le créez, ou vous pouvez le modifier ultérieurement.

2. Pour trier une réponse à une requête dans la console, choisissez l'attribut pour trier la réponse dans le menu Trier. Seuls les attributs marqués comme triables lors de la configuration de la source de données apparaissent dans la liste.

Réduction/extension des résultats de requête

> **Note**
>
> La prise en charge des fonctionnalités varie en fonction du type d'index et de l'API de recherche utilisés. Pour savoir si cette fonctionnalité est prise en charge pour le type d'index et l'API de recherche que vous utilisez, consultez la section Types d'index.

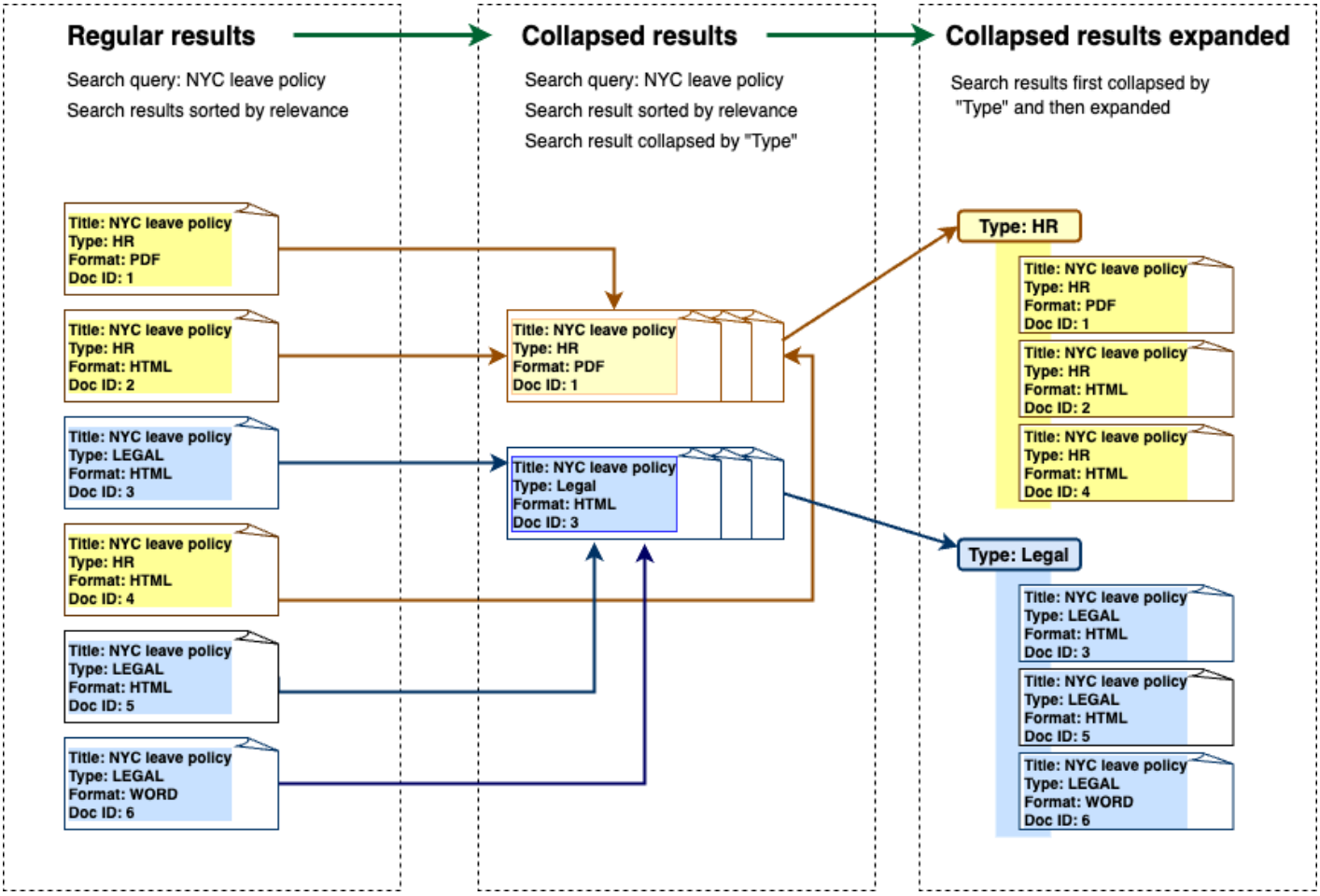
Lorsque vous vous connectez Amazon Kendra à vos données, il explore les attributs des métadonnées des documents (tels que `_document_title_created_at`, et) et utilise `_document_id` ces attributs ou champs pour fournir des fonctionnalités de recherche avancées au moment des requêtes.

Amazon Kendra La fonctionnalité Réduire et développer les résultats de requête vous permet de regrouper les résultats de recherche à l'aide d'un attribut de document commun et de les afficher (réduits ou partiellement développés) sous un document principal désigné.

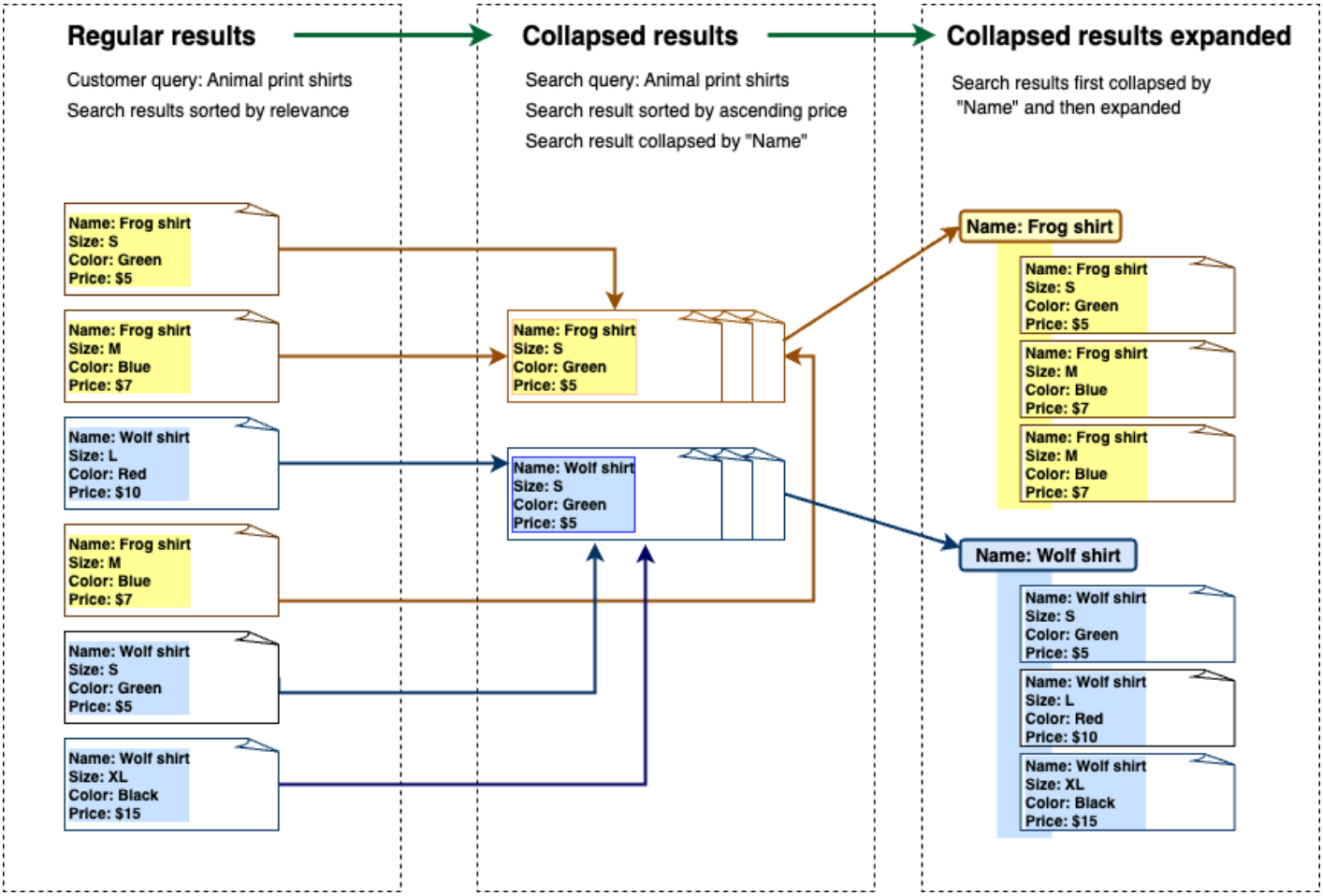
> **Note**
>
> La fonctionnalité de réduction et d'extension des résultats de requête n'est actuellement disponible que via l'Amazon Kendra API.

Cela est utile dans les situations de recherche suivantes :

- Plusieurs versions du contenu existent dans les documents de votre index. Lorsque votre utilisateur final interroge l'index, vous souhaitez qu'il voie la version la plus pertinente du document contenant des doublonshidden/collapsed. For example, if your index contains multiple versions of a document named "NYC leave policy" you can choose to collapse the documents for the specific groups "HR" and "Legal" using the "Type" attribute/field.



- Votre index contient plusieurs documents contenant des informations uniques sur un type d'article ou d'objet, comme un inventaire de produits, par exemple. Pour saisir et trier facilement les informations sur les articles, vous souhaitez que les utilisateurs finaux accèdent à tous les documents liés par un élément ou un objet sous la forme d'un seul résultat de recherche. Dans l'exemple ci-dessous, une recherche effectuée par un client sur « chemises à imprimé animalier » renvoie des résultats groupés par nom et triés par ordre de prix croissant.



Réduction des résultats

Pour regrouper des documents similaires ou connexes, vous devez spécifier l'attribut sur lequel vous souhaitez réduire (par exemple, vous pouvez réduire/regrouper des documents par`_category`). Pour ce faire, appelez l'API Query et utilisez l'CollapseConfigurationobjet pour spécifier l'objet `DocumentAttributeKey` à réduire. Les `DocumentAttributeKey` commandes sur lesquelles les résultats de recherche de champs seront réduits. Les champs clés d'attribut pris en charge incluent `String` et`Number`. `String list`et le `Date` type ne sont pas pris en charge.

Choix d'un document principal par ordre de tri

Pour configurer le document principal à afficher pour un groupe réduit, utilisez le `SortingConfigurations` paramètre ci-dessous CollapseConfiguration. Par exemple, pour obtenir la version la plus récente d'un document, vous devez trier chaque groupe réduit par`_version`. Vous pouvez en spécifier jusqu'à 3 attributes/fields to sort by and a sort order for each attribute/field en

utilisant`SortingConfigurations`. Vous pouvez demander une augmentation du quota pour le nombre d'attributs de tri.

Par défaut, Amazon Kendra trie les réponses aux requêtes en fonction du score de pertinence qu'il détermine pour chaque résultat de la réponse. Pour modifier l'ordre de tri par défaut, rendez les attributs du document triables, puis configurez Amazon Kendra pour utiliser ces attributs pour trier les réponses. Pour plus d'informations, consultez la section Tri des réponses.

Document manquant : stratégie clé

Si votre document ne possède pas de valeur d'attribut de réduction, il Amazon Kendra propose trois options de personnalisation :

- Choisissez d'inclure `COLLAPSE` tous les documents contenant des valeurs nulles ou manquantes dans un seul groupe. Il s'agit de la configuration par défaut.
- Choisissez les `IGNORE` documents dont les valeurs sont nulles ou manquantes. Les documents ignorés n'apparaîtront pas dans les résultats de la requête.
- Choisissez de `EXPAND` placer chaque document contenant une valeur nulle ou manquante dans un groupe distinct.

Élargir les résultats

Vous pouvez choisir d'étendre les groupes de résultats de recherche réduits en utilisant le `Expand` paramètre de l'CollapseConfigurationobjet. Les résultats étendus conservent le même ordre de tri que celui utilisé pour sélectionner le document principal du groupe.

Pour configurer le nombre de groupes de résultats de recherche réduits à développer, vous utilisez le `MaxResultItemstoExpand` paramètre de l'ExpandConfigurationobjet. Si vous définissez cette valeur sur 10, par exemple, seuls les 10 premiers groupes de résultats sur 100 disposeront de fonctionnalités étendues.

Pour configurer le nombre de résultats étendus à afficher par document principal réduit, utilisez le `MaxExpandResultsPerItem` paramètre. Par exemple, si vous définissez cette valeur sur 3, au maximum 3 résultats par groupe réduit seront affichés.

Interactions avec d'autres Amazon Kendra fonctionnalités

- La réduction et l'extension des résultats ne modifient pas le nombre de facettes et n'ont aucune incidence sur le nombre total de résultats affichés.

- Amazon Kendra les résultats de recherche présentés ne seront pas réduits même s'ils ont la même valeur de champ que le champ de réduction que vous configurez.
- La réduction et l'extension des résultats ne s'appliquent qu'aux résultats de type`DOCUMENT`.

Optimisation de la pertinence des recherches

> **Note**
>
> La prise en charge des fonctionnalités varie en fonction du type d'index et de l'API de recherche utilisés. Pour savoir si cette fonctionnalité est prise en charge pour le type d'index et l'API de recherche que vous utilisez, consultez la section Types d'index.

Amazon Kendra les requêtes produisent des résultats de recherche classés en fonction de leur pertinence. Les champs ou attributs consultables de l'index contribuent tous à ce classement.

Vous pouvez modifier l'effet d'un champ ou d'un attribut sur la pertinence de la recherche en ajustant la pertinence. Le réglage de la pertinence de la recherche peut être effectué manuellement au niveau de l'index, où vous définissez les configurations de réglage de votre index, ou au niveau de la requête en remplaçant les configurations définies au niveau de l'index.

Lorsque vous utilisez le réglage de la pertinence, le résultat est amélioré dans la réponse lorsque la requête inclut des termes correspondant au champ ou à l'attribut. Vous spécifiez également le niveau d'augmentation que le document reçoit en cas de correspondance. Le réglage de la pertinence n' Amazon Kendra entraîne pas l'inclusion d'un document dans la réponse à la requête, il ne s'agit que d'un des facteurs Amazon Kendra utilisés pour déterminer la pertinence d'un document.

Vous pouvez améliorer des champs ou des attributs spécifiques dans votre index afin d'attribuer plus d'importance à des réponses spécifiques. Par exemple, lorsque quelqu'un recherche « When is re:Invent ? » vous pourriez améliorer la pertinence de la fraîcheur des documents `_last_update_at` sur le terrain. Ou, dans un index de rapports de recherche, vous pouvez renforcer une source de données spécifique dans le champ « source ».

Vous pouvez également améliorer les documents en fonction des votes ou du nombre de vues, ce qui est courant dans les forums et autres bases de connaissances d'assistance. Vous pouvez combiner les boosters, par exemple pour améliorer les documents les plus consultés et les plus récents.

Vous définissez le niveau de boost qu'un document reçoit à l'aide du `Importance` paramètre. Plus le niveau est élevé`Importance`, plus le champ ou l'attribut renforce la pertinence d'un document. Lorsque vous réglez votre index ou que vous réglez au niveau de la requête, augmentez la valeur du `Importance` paramètre par petits incréments jusqu'à obtenir l'effet souhaité. Pour déterminer si vous

améliorez les résultats de recherche, effectuez la recherche et comparez les résultats aux requêtes précédentes.

Vous pouvez spécifier des attributs de date, de nombre ou de chaîne pour ajuster un index ou régler au niveau de la requête. Vous pouvez régler les champs ou les attributs de ce type `StringList` uniquement au niveau de l'index. Chaque champ ou attribut possède des critères spécifiques pour améliorer un résultat.

- Champs ou attributs de date : il existe trois critères spécifiques pour les champs de date`Duration`, `Freshness` et`RankOrder`.
  - `Duration`définit la période à laquelle le boost s'applique. Par exemple, si vous définissez la durée sur 86 400 secondes (c'est-à-dire un jour), le boost commence à diminuer au bout d'un jour. Plus l'importance est élevée, plus l'effet boost diminue rapidement.
  - `Freshness`détermine dans quelle mesure un document est récent lorsqu'il est appliqué à un champ ou à un attribut. Si vous postulez dans le champ correspondant `Freshness` à la date de création ou à la date de dernière mise à jour, un document créé récemment ou mis à jour est considéré comme « plus récent » qu'un document plus ancien. Par exemple, si le document 1 a été créé le 14 novembre et que le document 2 a été créé le 5 novembre, le document 1 est « plus récent » que le document 2. Et si le document 1 a été mis à jour pour la dernière fois le 14 novembre et que le document 2 a été mis à jour pour la dernière fois le 20 novembre, le document 2 est « plus récent » que le document 1. Plus le document est frais, plus cette amélioration est appliquée. Vous ne pouvez avoir qu'un seul `Freshness` champ dans votre index.
  - `RankOrder`applique le boost par ordre croissant ou décroissant. Si vous le spécifiez`ASCENDING`, les dates ultérieures ont priorité. Si vous le spécifiez`DESCENDING`, les dates antérieures ont priorité.
- Champs numériques ou attributs : pour les champs numériques ou les attributs, vous pouvez spécifier l'ordre de classement Amazon Kendra à utiliser pour déterminer la pertinence du champ ou de l'attribut. Si vous le spécifiez`ASCENDING`, les nombres les plus élevés ont la priorité. Si vous le spécifiez`DESCENDING`, les nombres inférieurs ont priorité.
- Champs ou attributs de chaîne : pour les champs de chaîne ou les attributs, vous pouvez créer des catégories d'un champ afin de donner une impulsion différente à chaque catégorie. Par exemple, si vous augmentez un champ ou un attribut appelé « Département », vous pouvez donner une valeur différente aux documents issus de la catégorie « RH » et aux documents de la catégorie « Juridique ». Vous pouvez améliorer un champ ou un attribut de ce type`String`. Vous pouvez augmenter `StringList` les champs uniquement au niveau de l'index.

Réglage de la pertinence au niveau de l'indice

Vous pouvez ajuster la pertinence d'un champ ou d'un attribut au niveau de l'index en utilisant soit la console pour définir les détails de l'index, soit l'UpdateIndexAPI.

L'exemple suivant définit le `_last_updated_at` champ comme `Freshness` champ d'un document.

```
"DocumentMetadataConfigurationUpdates" : [
    {
        "Name": "_last_updated_at",
        "Type": "DATE_VALUE",
        "Relevance": {
            "Freshness": TRUE,
            "Importance": 2
        }
    }
]
```

L'exemple suivant applique une importance différente aux différentes catégories du champ « département ».

```
"DocumentMetadataConfigurationUpdates" : [
    {
        "Name": "department",
        "Type": "STRING_VALUE",
        "Relevance": {
            "Importance": 2,
            "ValueImportanceMap": {
                "HR": 3,
                "Legal": 1
            }
        }
    }
]
```

Réglage de la pertinence au niveau de la requête

Vous pouvez ajuster la pertinence d'un champ ou d'un attribut au niveau de la requête à l'aide de l'API Query.

Le réglage de la pertinence au niveau de la requête n'est pas pris en charge dans la console.

Le réglage au niveau de la requête peut accélérer le processus de test du réglage de la pertinence, car il n'est pas nécessaire de mettre à jour manuellement les configurations de réglage dans l'index pour chaque test. Vous pouvez ajuster la pertinence d'un document en transmettant les configurations de réglage dans la requête. Ensuite, vous pouvez voir les différents résultats que vous obtenez avec les différentes configurations. Une configuration transmise dans la requête remplace la configuration définie au niveau de l'index.

L'exemple suivant remplace l'importance accordée au champ « département » et à chaque catégorie de département définie au niveau de l'index, comme indiqué dans l'exemple ci-dessus. Lorsqu'un utilisateur saisit sa requête de recherche, le champ « département » a un certain niveau d'importance et le service juridique a plus d'importance que le service des ressources humaines.

```
"DocumentRelevanceOverrideConfigurations" : [
    {
        "Name": "department",
        "Type": "STRING_VALUE",
        "Relevance": {
            "Importance": 2,
            "ValueImportanceMap": {
                "HR": 2,
                "Legal": 8
            }
        }
    }
]
```

Obtenir des informations grâce à l'analyse des recherches

> **Note**
>
> La prise en charge des fonctionnalités varie en fonction du type d'index et de l'API de recherche utilisés. Pour savoir si cette fonctionnalité est prise en charge pour le type d'index et l'API de recherche que vous utilisez, consultez la section Types d'index.

Vous pouvez utiliser l'analyse de Amazon Kendra recherche pour savoir comment votre application de recherche aide ou non vos utilisateurs à trouver des informations.

Amazon Kendra Les analyses fournissent un aperçu de la manière dont vos utilisateurs interagissent avec votre application de recherche et de l'efficacité de la configuration de votre application de recherche. Vous pouvez consulter les données des métriques à l'aide de l'GetSnapshotsAPI ou en sélectionnant Analytics dans le panneau de navigation de la console.

Vous pouvez afficher les données générées par `GetSnapshots` votre propre tableau de bord personnalisé. Vous pouvez également utiliser le tableau de bord des métriques fourni dans la console, qui inclut des graphiques visuels. Grâce à un tableau de bord visuel, vous pouvez identifier les tendances ou les modèles du comportement des utilisateurs au fil du temps ou détecter des problèmes liés à la configuration de votre application de recherche. Par exemple, un graphique linéaire qui montre un nombre constant de requêtes par jour et une augmentation constante peut indiquer une adoption et une utilisation accrues. D'un autre côté, une chute brutale peut indiquer qu'un problème doit être étudié.

Vous pouvez utiliser les métriques pour établir des liens entre différents points de données afin de résoudre les problèmes liés à la manière dont vos utilisateurs recherchent des informations ou découvrent des opportunités commerciales. Par exemple, le document « Comment fonctionne l'IA ? » est le document le plus cliqué dans les résultats de recherche, et la requête la plus recherchée est « Comment fonctionne le machine learning ? ». Cela vous renseigne sur les termes et la langue préférés utilisés par vos utilisateurs. Vous pouvez intégrer ces termes dans vos documents ou utiliser des synonymes personnalisés pour les rendre plus consultables par vos utilisateurs.

Métriques pour la recherche

Il existe 10 indicateurs permettant d'analyser les performances de votre application de recherche ou les informations recherchées par vos utilisateurs. Pour récupérer les données métriques, vous

devez spécifier le nom de chaîne des données métriques que vous souhaitez récupérer lorsque vous appelez`GetSnapshots`.

Vous devez également fournir un intervalle de temps ou une fenêtre de temps pour afficher les données des métriques. L'intervalle de temps utilise le fuseau horaire de votre index. Vous pouvez consulter les données dans les fenêtres temporelles suivantes :

- `THIS_WEEK`: La semaine en cours, commençant le dimanche et se terminant la veille de la date en cours.
- `ONE_WEEK_AGO`: La semaine précédente, commençant le dimanche et se terminant le samedi suivant.
- `TWO_WEEKS_AGO`: La semaine précédant la semaine précédente, commençant le dimanche et se terminant le samedi suivant.
- `THIS_MONTH`: Le mois en cours, commençant le premier jour du mois et se terminant la veille de la date en cours.
- `ONE_MONTH_AGO`: Le mois précédent, commençant le premier jour du mois et se terminant le dernier jour du mois.
- `TWO_MONTHS_AGO`: Le mois précédant le mois précédent, commençant le premier jour du mois et se terminant le dernier jour du mois.

Dans la console, les fenêtres horaires prises en charge sont Cette semaine, Semaine précédente, Ce mois, Mois précédent.

Taux de clics

Proportion de requêtes qui mènent à un clic vers un document dans les résultats de recherche. Cela vous permet de savoir si la configuration de votre application de recherche aide les utilisateurs à trouver les informations pertinentes pour leurs requêtes. Pour les requêtes qui renvoient des réponses instantanées, les utilisateurs n'ont peut-être pas besoin de cliquer sur un document pour obtenir plus d'informations. Pour de plus amples informations, veuillez consulter the section called “Taux de réponse instantané”. Vous devez appeler SubmitFeedbackpour vous assurer que les commentaires reçus lors des clics sont collectés.

Pour récupérer des données sur le taux de clics à l'aide de l'`GetSnapshots`API, spécifiez le `metricType` as. `AGG_QUERY_DOC_METRICS` Vous pouvez également consulter cette métrique dans la console en sélectionnant Analytics dans le panneau de navigation.

Taux de clics nul

Proportion de requêtes n'entraînant aucun clic dans les résultats de recherche. Cela vous permet de comprendre les lacunes de votre contenu en fournissant des résultats de recherche non pertinents. Pour les requêtes qui renvoient des réponses instantanées, les utilisateurs n'ont peut-être pas besoin de cliquer sur un document pour obtenir plus d'informations. Pour de plus amples informations, veuillez consulter the section called “Taux de réponse instantané”. En outre, vos paramètres de recherche, tels que le réglage des configurations, peuvent avoir un impact sur la manière dont les documents sont renvoyés dans les résultats de recherche.

Pour récupérer des données avec un taux de clic nul à l'aide de l'`GetSnapshots`API, spécifiez le `metricType` as`AGG_QUERY_DOC_METRICS`. Vous pouvez également consulter cette métrique dans la console en sélectionnant Analytics dans le panneau de navigation.

Taux de résultats de recherche nul

Proportion de requêtes n'aboutissant à aucun résultat de recherche. Cela vous permet de comprendre les lacunes de votre contenu en ne fournissant aucun résultat de recherche pertinent.

Pour récupérer des données sur un taux de résultats de recherche nul à l'aide de l'`GetSnapshots`API, spécifiez le `metricType` as`AGG_QUERY_DOC_METRICS`. Vous pouvez également consulter cette métrique dans la console en sélectionnant Analytics dans le panneau de navigation.

Taux de réponse instantané

Proportion de requêtes ayant reçu une réponse instantanée ou une FAQ renvoyées. Cela vous aide à comprendre le rôle des réponses instantanées dans la fourniture d'informations.

Pour récupérer des données sur le taux de réponse instantané à l'aide de l'`GetSnapshots`API, spécifiez le `metricType` as`AGG_QUERY_DOC_METRICS`. Vous pouvez également consulter cette métrique dans la console en sélectionnant Analytics dans le panneau de navigation.

Requêtes les plus fréquentes

Les 100 requêtes les plus recherchées par vos utilisateurs. Cela vous permet de comprendre quelles sont les requêtes les plus populaires et le type d'informations qui intéressent le plus vos utilisateurs.

Les indicateurs incluent le nombre de fois où la requête a fait l'objet d'une recherche, la proportion de clics sur un document, la proportion de clics non cliqués sur un document, la profondeur moyenne

des clics dans les résultats de recherche pour la requête, la proportion de réponses instantanées à la requête et le niveau de confiance moyen pour les 10 premiers résultats de recherche d'une requête.

Pour récupérer des données sur les principales requêtes à l'aide de l'`GetSnapshots`API, spécifiez le `metricType` as`QUERIES_BY_COUNT`. Vous pouvez également consulter cette métrique dans la console en sélectionnant Analytics dans le panneau de navigation de la console, puis en sélectionnant les requêtes les plus fréquentes sous Listes de requêtes.

Requêtes les plus fréquentes sans aucun clic

Les 100 requêtes les plus fréquentes qui n'ont généré aucun clic dans les résultats de recherche. Cela vous permet de comprendre les lacunes dans votre contenu, en cas de manque de documents pertinents pour certaines requêtes ou lorsque la configuration de votre application de recherche renvoie des résultats de recherche non pertinents. Pour les requêtes qui renvoient des réponses instantanées, les utilisateurs n'ont peut-être pas besoin de cliquer sur un document pour obtenir plus d'informations. Pour de plus amples informations, veuillez consulter the section called “Taux de réponse instantané”.

Les indicateurs incluent le nombre de fois où la requête n'a généré aucun clic, la proportion de zéro clic pour la requête, la proportion de réponses instantanées pour la requête et le niveau de confiance moyen pour les 10 premiers résultats de recherche d'une requête.

Pour récupérer des données sur les principales requêtes sans aucun clic à l'aide de l'`GetSnapshots`API, spécifiez le `metricType` as`QUERIES_BY_ZERO_CLICK_RATE`. Vous pouvez également consulter cette métrique dans la console en sélectionnant Analytics dans le panneau de navigation de la console, puis en sélectionnant les requêtes à zéro clic les plus fréquentes sous Listes de requêtes.

Requêtes les plus fréquentes sans aucun résultat

Les 100 requêtes les plus fréquentes qui n'ont donné aucun résultat de recherche. Cela vous permet de comprendre toute lacune dans votre contenu, lorsqu'aucun document ne correspond à certaines requêtes. Vos utilisateurs peuvent également utiliser des termes spécialisés susceptibles de ne pas aboutir à des résultats de recherche, ce qui vous invite à créer des synonymes personnalisés pour gérer cette situation.

Les mesures incluent le nombre de fois où la requête n'aboutit à aucun résultat de recherche, la proportion de zéro résultat de recherche pour la requête et la proportion de fois où la requête est recherchée par rapport à toutes les requêtes.

Pour récupérer des données sur les requêtes les plus fréquentes sans aucun résultat de recherche à l'aide de l'`GetSnapshots`API, spécifiez le `metricType` as`QUERIES_BY_ZERO_RESULT_RATE`. Vous pouvez également consulter cette métrique dans la console en sélectionnant Analytics dans le panneau de navigation de la console, puis en sélectionnant les requêtes ayant obtenu les meilleurs résultats sous Listes de requêtes.

Le haut de la page a cliqué sur les documents

Les 100 documents les plus consultés dans les résultats de recherche. Cela vous permet de comprendre quels documents ou quels résultats de recherche sont les plus pertinents pour vos utilisateurs lorsqu'ils demandent des informations.

Les indicateurs incluent le nombre de fois où l'on clique sur le document, le nombre de likes qu'un document reçoit de la part de vos utilisateurs (pouce levé), le nombre de dégoûts qu'un document reçoit de la part de vos utilisateurs (pouce vers le bas).

Pour récupérer les données sur les documents sur lesquels vous avez cliqué en haut à l'aide de l'`GetSnapshots`API, spécifiez le `metricType` as`DOCS_BY_CLICK_COUNT`. Vous pouvez également consulter cette métrique dans la console en sélectionnant Analytics dans le panneau de navigation de la console, puis en sélectionnant les documents les plus cliqués sous Listes de requêtes.

Nombre total de requêtes

Le nombre total de requêtes recherchées par vos utilisateurs. Cela vous permet de comprendre dans quelle mesure vos utilisateurs interagissent avec votre application de recherche.

Pour récupérer des données sur le nombre total de requêtes à l'aide de l'`GetSnapshots`API, spécifiez le `metricType` as`AGG_QUERY_DOC_METRICS`. Vous pouvez également consulter cette métrique dans la console en sélectionnant Analytics dans le panneau de navigation.

Total des documents

Le nombre total de documents dans votre index. Cela vous permet de comparer la taille de votre index au nombre total de requêtes afin de vérifier s'il existe un nombre de documents adapté au volume de requêtes.

Pour récupérer des données sur l'ensemble des documents à l'aide de l'`GetSnapshots`API, spécifiez le `metricType` as`AGG_QUERY_DOC_METRICS`. Vous pouvez également consulter cette métrique dans la console en sélectionnant Analytics dans le panneau de navigation.

Exemple de récupération de données métriques

Le code suivant est un exemple de récupération de données sur les principales requêtes du mois précédent.

Console

Pour récupérer les requêtes les plus fréquentes du mois précédent

1. Dans le volet de navigation de gauche, sous Index, sélectionnez votre index, puis Analytics.
2. Sur la page Analytics, sélectionnez le bouton Cette semaine pour modifier la fenêtre temporelle de récupération des données en Mois précédent.
3. Sur la page Analytics, sous Listes de requêtes, sélectionnez Principales requêtes.

CLI

Pour récupérer les requêtes les plus fréquentes du mois précédent

```
aws kendra get-snapshots \
--index-id index-id \
--interval "ONE_MONTH_AGO" \
--metric-type "QUERIES_BY_COUNT"
```

Python

Pour récupérer les requêtes les plus fréquentes du mois précédent

```
import boto3

kendra = boto3.client("kendra")

index_id = "index-id"
interval = "ONE_MONTH_AGO"
metric_type = "QUERIES_BY_COUNT"

snapshots_response = kendra.get_snapshots(
    IndexId = index_id,
    Interval = interval,
    MetricType = metric_type
)
```

```
print("Top queries data: " + snapshots_response["snapshotsData"])
```

Java

Pour récupérer les requêtes les plus fréquentes du mois précédent

```
package com.amazonaws.kendra;

import software.amazon.awssdk.services.kendra.KendraClient;
import software.amazon.awssdk.services.kendra.model.GetSnapshotsRequest;
import software.amazon.awssdk.services.kendra.model.GetSnapshotsResponse;

public class TopQueriesExample {
    public static void main(String[] args) {
        KendraClient kendra = KendraClient.builder().build();

        String indexId = "indexID";
        String interval = "ONE_MONTH_AGO";
        String metricType = "QUERIES_BY_COUNT";

        GetSnapshotsRequest getSnapshotsRequest = GetSnapshotsRequest
            .builder()
            .indexId(indexId)
            .interval(interval)
            .metricType(metricType)
            .build();

        GetSnapshotsResponse getSnapshotsResponse =
 kendra.getSnapshots(GetSnapshotsRequest);

        System.out.println(String.format("Top queries data: ",
 getSnapshotsResponse.snapshotsData()))
```

Des indicateurs aux informations exploitables

Les informations exploitables sont des informations pertinentes extraites de données brutes et utilisées pour guider vos actions ou vos décisions. Pour extraire du sens des indicateurs et les utiliser pour générer des informations exploitables, il est important non seulement d'examiner les indicateurs de manière isolée, mais également d'établir des liens entre les indicateurs.

Par exemple, la première requête sans aucun clic est « Quelles régions sont actuellement disponibles ? ». Cependant, il a également un taux de réponse instantanée de 100 %. Cela suggère que vos utilisateurs reçoivent la réponse à cette question sans avoir à cliquer sur un résultat de recherche ou un document fournissant des informations sur les régions disponibles. Si vous considérez le zéro clic uniquement, vous n'obtiendrez pas l'histoire complète et vous pourriez tirer de mauvaises conclusions quant au succès de la configuration de votre application de recherche dans le traitement de cette requête.

La découverte d'une opportunité commerciale est un autre exemple d'information exploitable. Les entreprises recherchent souvent des opportunités pour développer leurs clients en analysant les indicateurs de recherche. Le document le plus cliqué est « Régions disponibles ». En outre, la plupart des requêtes les plus recherchées concernent des questions sur la disponibilité des produits dans la région océanique, avec un taux de réponse instantanée de 100 % et un taux de clics élevé pour obtenir plus d'informations sur les régions disponibles dans le cadre de la réponse. Cela suggère que votre produit ou service suscite de l'intérêt et de la demande dans cette région.

Visualisation et génération de rapports sur les analyses de recherche

Cinq indicateurs incluent des données sur les tendances que vous pouvez visualiser et rechercher des tendances ou des modèles au fil du temps. Si vous utilisez la console, des graphiques des données de tendances sont fournis. Si vous utilisez le APIs, vous pouvez récupérer les données de tendances pour créer vos propres graphiques ou visualisations. La plupart des graphiques de la console tracent les points de données quotidiens sur la période que vous avez choisie.

La console fournit un tableau de bord des indicateurs dans lequel vous pouvez sélectionner un graphique et la liste des principaux éléments que vous souhaitez consulter. Vous pouvez exporter les statistiques affichées sur votre tableau de bord au format CSV en sélectionnant Exporter sur la page d'accueil d'Analytics. Vous pouvez inclure ces rapports dans vos documents commerciaux ou vos présentations.

Vous pouvez visualiser les indicateurs suivants :

Graphique du nombre total de requêtes

Un graphique linéaire du nombre de requêtes émises par jour. Le graphique vous aide à visualiser les modèles d'engagement quotidien des utilisateurs. Parmi les exemples, citons une augmentation

ou une diminution constante de l'engagement des utilisateurs, ou une chute drastique à 0 requêtes en raison d'un crash de votre application de recherche ou de problèmes avec votre site Web.

Si vous utilisez l'API, vous pouvez récupérer ces données en spécifiant`TREND_QUERY_DOC_METRICS`. Vous pouvez utiliser les données pour créer vos propres graphiques ou utiliser les graphiques fournis dans la console.

Graphique du taux de clics

Un graphique linéaire des proportions de clics par jour. Le graphique vous aide à visualiser les tendances du taux de clics quotidien. Parmi les exemples, citons une augmentation ou une diminution constante du taux de clics, ou une diminution du nombre de réponses instantanées susceptible d'influencer une augmentation du taux de clics.

Si vous utilisez l'API, vous pouvez récupérer ces données en spécifiant`TREND_QUERY_DOC_METRICS`. Vous pouvez utiliser les données pour créer vos propres graphiques ou utiliser les graphiques fournis dans la console.

Graphique du taux de clic nul

Un graphique linéaire de la proportion de zéro clic par jour. Le graphique vous aide à visualiser les modèles de taux de clics quotidiens nuls. Parmi les exemples, citons une augmentation ou une diminution constante du taux de clic nul, ou une augmentation du nombre de réponses instantanées susceptible d'influencer une augmentation du taux de zéro clic.

Si vous utilisez l'API, vous pouvez récupérer ces données en spécifiant`TREND_QUERY_DOC_METRICS`. Vous pouvez utiliser les données pour créer vos propres graphiques ou utiliser les graphiques fournis dans la console.

Graphique du taux de résultats de recherche nuls

Un graphique linéaire de la proportion de zéro résultat de recherche par jour. Le graphique vous aide à visualiser les tendances du taux quotidien de zéro résultat de recherche. Parmi les exemples, citons une augmentation ou une diminution constante du taux de résultats de recherche nuls, ou une forte diminution du nombre de documents dans votre index, susceptible d'influencer une augmentation du taux de zéro résultat de recherche.

Si vous utilisez l'API, vous pouvez récupérer ces données en spécifiant`TREND_QUERY_DOC_METRICS`. Vous pouvez utiliser les données pour créer vos propres graphiques ou utiliser les graphiques fournis dans la console.

Graphique du taux de réponse instantané

Un graphique linéaire de la proportion de requêtes ayant reçu une réponse instantanée ou une FAQ renvoyées. Le graphique vous aide à visualiser les tendances du taux de réponse instantanée quotidien. Parmi les exemples, citons l'augmentation ou la diminution constante du nombre de requêtes de type question-réponse, ou la diminution du nombre de clics susceptibles d'influencer une augmentation du nombre de réponses instantanées.

Si vous utilisez l'API, vous pouvez récupérer ces données en spécifiant`TREND_QUERY_DOC_METRICS`. Vous pouvez utiliser les données pour créer vos propres graphiques ou utiliser les graphiques fournis dans la console.

Soumission de commentaires pour un apprentissage progressif

> **Note**
>
> La prise en charge des fonctionnalités varie en fonction du type d'index et de l'API de recherche utilisés. Pour savoir si cette fonctionnalité est prise en charge pour le type d'index et l'API de recherche que vous utilisez, consultez la section Types d'index.

Amazon Kendra utilise l'apprentissage progressif pour améliorer les résultats de recherche. En utilisant les commentaires issus des requêtes, l'apprentissage progressif améliore les algorithmes de classement et optimise les résultats de recherche pour une plus grande précision.

Supposons, par exemple, que vos utilisateurs recherchent l'expression « prestations de santé ». Si les utilisateurs choisissent régulièrement le deuxième résultat de la liste, ce résultat Amazon Kendra passe au fil du temps à la première place. Le boost diminue avec le temps, donc si les utilisateurs arrêtent de sélectionner un résultat, le Amazon Kendra suppriment finalement et affichent un autre résultat plus populaire à la place. Cela permet de Amazon Kendra hiérarchiser les résultats en fonction de leur pertinence, de leur âge et de leur contenu.

L'apprentissage progressif est activé pour tous les index et pour tous les types de documents pris en charge.

Amazon Kendra commence à apprendre dès que vous fournissez des commentaires, mais les résultats des commentaires peuvent prendre plus de 24 heures. Amazon Kendra propose trois méthodes pour envoyer des commentaires : la AWS console, une JavaScript bibliothèque que vous pouvez inclure sur votre page de résultats de recherche et une API que vous pouvez utiliser.

Amazon Kendra accepte deux types de commentaires des utilisateurs :

- Clics : informations sur les résultats de requête choisis par l'utilisateur. Les commentaires incluent l'ID du résultat et l'horodatage Unix de la date et de l'heure auxquelles le résultat de recherche a été choisi.

  Pour envoyer des commentaires sur les clics, votre application doit collecter des informations sur les clics provenant des activités de vos utilisateurs, puis les envoyer à Amazon Kendra. Vous

pouvez collecter des informations sur les clics à l'aide de la console, de la JavaScript bibliothèque et de l' Amazon Kendra API.

- Pertinence : informations relatives à la pertinence d'un résultat de recherche, généralement fournies par l'utilisateur. Le feedback contient l'identifiant du résultat et un indicateur de pertinence (`RELEVANTouNOT_RELEVANT`). L'utilisateur détermine les informations de pertinence.

  Pour envoyer des commentaires sur la pertinence, votre application doit fournir un mécanisme de feedback qui permet à l'utilisateur de choisir la pertinence appropriée pour le résultat d'une requête, puis de soumettre ces informations à Amazon Kendra. Vous ne pouvez collecter des informations de pertinence qu'avec la console et l' Amazon Kendra API.

Les commentaires sont utilisés lorsque l'index est actif. Les commentaires n'affectent que l'index auquel ils sont soumis, ils ne peuvent pas être utilisés entre les index ou pour différents comptes.

Vous devez fournir un contexte utilisateur supplémentaire lorsque vous interrogez votre Amazon Kendra index. Lorsque vous fournissez un contexte utilisateur, Amazon Kendra il est en mesure de savoir si les commentaires sont fournis par un seul utilisateur ou par plusieurs utilisateurs et d'ajuster les résultats de recherche en conséquence.

Lorsque vous fournissez un contexte utilisateur, les commentaires relatifs à la requête sont associés à l'utilisateur spécifique indiqué dans le contexte. Si vous ne spécifiez pas le contexte utilisateur, vous pouvez fournir un identifiant de visiteur utilisé pour regrouper et agréger les requêtes.

Si vous ne fournissez pas de contexte utilisateur ou d'identifiant de visiteur, les commentaires sont anonymes et agrégés avec d'autres commentaires anonymes.

Le code suivant montre comment inclure le contexte utilisateur sous forme de jeton ou d'identifiant de visiteur.

```
response = kendra.query(
    QueryText = query,
    IndexId = index,
    UserToken = {
        Token = "token"
    })

    OR

    response = kendra.query(
    QueryText = query,
```

```
    IndexId = index,
    VisitorId = "visitor-id")
```

Pour les applications Web, vous pouvez utiliser des cookies, des localisations ou des utilisateurs de navigateur pour générer un identifiant de visiteur pour chaque utilisateur.

Pour les requêtes principales, le plus grand volume de requêtes, la fourniture de commentaires en un clic fournit suffisamment d'informations pour améliorer la précision globale. Pour les requêtes finales, celles qui sont rares, les experts en la matière doivent soumettre des commentaires pertinents et non pertinents afin d'améliorer la précision de ces requêtes.

Outre la console, vous pouvez utiliser l'une des deux méthodes suivantes : une JavaScript bibliothèque ou l'SubmitFeedbackAPI. Vous ne devez utiliser qu'une seule méthode pour recueillir des commentaires. Pour de meilleurs résultats, vous devez envoyer vos commentaires dans les 24 heures suivant la demande.

Rubriques

- Utiliser la Amazon Kendra JavaScript bibliothèque pour envoyer des commentaires
- Utiliser l' Amazon Kendra API pour envoyer des commentaires

Utiliser la Amazon Kendra JavaScript bibliothèque pour envoyer des commentaires

Amazon Kendra fournit une JavaScript bibliothèque que vous pouvez utiliser pour ajouter des commentaires sur les clics à votre page de résultats de recherche. Pour utiliser la bibliothèque, vous insérez une balise de script dans votre code client qui affiche le résultat de la recherche, puis vous ajoutez des informations à chacun des liens de documents de votre liste de résultats. Lorsqu'un utilisateur choisit un lien pour afficher un document, les informations relatives aux clics sont envoyées à Amazon Kendra.

La bibliothèque fonctionne avec les navigateurs compatibles avec JavaScript la version ES6/ES2015.

Étape 1 : insérez une balise de script dans votre application Amazon Kendra de recherche

Dans votre code client qui affiche les résultats Amazon Kendra de recherche, insérez une <script>balise et ajoutez une référence à la JavaScript bibliothèque :

```
<script>
 (function(w, d, s, c, g, n) {
   if(!w[n]) {
     w[n] = w[n] || function () {
           (w[n].q = w[n].q || []).push(arguments);
     }
     w[n].st = new Date().getTime();
     w[n].ep = g;
     var e = document.createElement(s),
         j = document.getElementsByTagName(s)[0];
     e.async = 1;
     e.src = c;
     e.type = 'module';
     j.parentNode.insertBefore(e, j);
   }
 })(window, document, 'script',
 'library download URL',
 'feedback endpoint',
 'kendraFeedback');
</script>
```

Le script télécharge de manière asynchrone la JavaScript bibliothèque depuis un CDN Amazon Kendra hébergé et initialise une variable globale appelée `kendraFeedback` qui vous permet de définir des paramètres facultatifs.

Remplacez *`library download URL`* et *`feedback endpoint`* par un identifiant du tableau suivant en fonction de la région qui héberge votre Amazon Kendra index.

Région	Télécharger le kit URL	Point final de feedback
us-east-1	https://d2zm0lpns956f8.cloudfront.net/ksf-v1.js	https://ujxwp5s92h.execute-api.us-east-1.amazonaws.com/prod/soumettre
us-east-2	https://d2crv7fufeg244.cloudfront.net/ksf-v1.js	https://i6h76zwzf3.execute-api.us-east-2.amazonaws.com/prod/soumettre

Région	Télécharger le kit URL	Point final de feedback
us-west-2	https://d2iezfpnpcoujy.cloudfront.net/ksf-v1.js	https://wg6nim909c.execute-api.us-west-2.amazonaws.com/prod/soumettre
ca-central-1	https://d1zbkfomowykaq.cloudfront.net/ksf-v1.js	https://budi8txevj.execute-api.ca-central-1.amazonaws.com/prod/soumettre
eu-west-1	https://d3gptlxtulu4us.cloudfront.net/ksf-v1.js	https://po2b11740b.execute-api.eu-west-1.amazonaws.com/prod/soumettre
ap-southeast-1	https://d1vvuam7g4taoe.cloudfront.net/ksf-v1.js	https://9je5uw7t5l.execute-api.ap-southeast-1.amazonaws.com/prod/soumettre
ap-southeast-2	https://dopqntoe6z0ce.cloudfront.net/ksf-v1.js	https://oovf4nvjj7.execute-api.ap-southeast-2.amazonaws.com/prod/soumettre
ap-south-1	https://d1ts9ouelsmk3g.cloudfront.net/ksf-v1.js	https://k1abnmd43b.execute-api.ap-south-1.amazonaws.com/prod/soumettre
ap-northeast-1	https://d3w0ybsa293kb4.cloudfront.net/ksf-v1.js	https://wg7rz0uzjh.execute-api.ap-northeast-1.amazonaws.com/prod/soumettre
eu-west-2	https://d1tsrujswld1d1.cloudfront.net/ksf-v1.js	https://qi7mct3x7f.execute-api.eu-west-2.amazonaws.com/prod/soumettre

Par exemple, si votre indice est situé dans l'est des États-Unis (Virginie du Nord), *library download URL* il est `https://d2zm0lpns956f8.cloudfront.net/ksf-v1.js` et *feedback endpoint* est`https://ujxwp5s92h.execute-api.us-east-1.amazonaws.com/prod/submit`.

Vous pouvez définir deux paramètres facultatifs pour la Amazon Kendra JavaScript bibliothèque :

- `disableCookies`— Par défaut, Amazon Kendra définit un cookie qui identifie l'utilisateur de manière unique. Réglez ce paramètre `true` sur pour désactiver le cookie.

  ```
  kendraFeedback('disableCookie', 'true | false');
  ```

  `searchDivClassName`— Par défaut, Amazon Kendra surveille tous les liens de votre page de résultats de recherche pour détecter les clics. Définissez ce paramètre sur un nom de `<div>` classe pour surveiller uniquement les liens de la classe spécifiée.

  ```
  kendraFeedback('searchDivClassName', 'class name');
  ```

Étape 2 : ajouter le jeton de commentaires aux résultats de recherche

Sur votre page de résultats, ajoutez un attribut HTML appelé `data-kendra-token` à la balise d'ancrage ou à la balise div du parent immédiat qui contient un lien vers le document depuis la réponse à la requête. Par exemple :

```
<a href="document location" data-kendra-token="feedback token value"></a>
OR
<div data-url="document location" data-kendra-token="feedback token value"></div>
```

Une réponse à une requête contient un jeton dans le `feedbackToken` champ. Le jeton identifie de manière unique la réponse si l'utilisateur la choisit. Attribuez la valeur du jeton à l'`data-kendra-token`attribut. La Amazon Kendra JavaScript bibliothèque recherche ce jeton lorsque l'utilisateur choisit le résultat et le soumet à un Amazon Kendra point de terminaison en tant que commentaire.

La Amazon Kendra JavaScript bibliothèque envoie uniquement le jeton de commentaire et d'autres métadonnées, telles que l'heure à laquelle le résultat a été choisi et un identifiant de visiteur unique.

Étape 3 : tester le script de feedback

Pour vous assurer que la JavaScript bibliothèque est correctement configurée et pour envoyer des commentaires au point de terminaison approprié, procédez comme suit. Cet exemple utilise le navigateur Chrome.

1. Ouvrez les outils de développement Web dans le navigateur. Sur Chrome, ouvrez le menu Chrome dans le coin supérieur droit du navigateur, choisissez Plus d'outils, puis sélectionnez Outils pour développeurs.
2. Assurez-vous qu'aucune erreur liée à la Amazon Kendra JavaScript bibliothèque ne se trouve dans l'onglet console.
3. Effectuez une recherche et choisissez n'importe quel résultat. Dans l'onglet Réseau des outils de développement. Vous devriez voir une demande envoyée au point de terminaison de feedback, le jeton correspondant au résultat et un statut 200 OK.

Utiliser l' Amazon Kendra API pour envoyer des commentaires

Pour utiliser l' Amazon Kendra API pour envoyer des commentaires sur les requêtes, utilisez l'SubmitFeedbackAPI. Pour identifier la requête, vous devez fournir l'ID d'index de l'index auquel la requête s'applique et l'ID de requête renvoyé dans la réponse de l'API Query.

L'exemple suivant montre comment envoyer des commentaires sur les clics et la pertinence à l'aide de l' Amazon Kendra API. Vous pouvez envoyer plusieurs ensembles de commentaires par le biais des `RelevanceFeedbackItems` tableaux `ClickFeedbackItems` et. Cet exemple envoie un seul clic et un seul commentaire de pertinence. La soumission de commentaires utilise l'heure actuelle.

Pour envoyer des commentaires concernant une recherche (AWS SDK)

1. Vous pouvez utiliser l'exemple de code suivant avec les valeurs requises :

   a. `index id`: l'ID de l'index auquel s'applique la requête.

   b. `query id`: la requête sur laquelle vous souhaitez fournir des commentaires.

   c. `result id`: l'ID du résultat de la requête sur lequel vous souhaitez fournir des commentaires. La réponse à la requête contient l'ID du résultat.

   d. `relevance value`—Soit `RELEVANT` (le résultat de la requête est pertinent) soit `NOT_RELEVANT` (le résultat de la requête n'est pas pertinent).

   Python

```
import boto3
import time

kendra = boto3.client("kendra")
```

```
# Provide the index ID
index_id = "index-id"
# Provide the query ID
query_id = "query-id"
# Provide the search result ID
result_id = "result-id"

# Configure the feedback item
feedback_item = {"ClickTime": int(time.time()),
    "ResultId":result_id}

# Configure the relevance value
relevance_value = "RELEVANT"
relevance_item = {"RelevanceValue": relevance_value,
    "ResultId": result_id
    }

response = kendra.submit_feedback(
    QueryId = query_id,
    IndexId = index_id,
    ClickFeedbackItems = [feedback_item],
    RelevanceFeedbackItems = [relevance_item]
)


print("Submitted feedback for query: " + query_id)
```

Java

```
package com.amazonaws.kendra;

import java.time.Instant;
import software.amazon.awssdk.services.kendra.KendraClient;
import software.amazon.awssdk.services.kendra.model.ClickFeedback;
import software.amazon.awssdk.services.kendra.model.RelevanceFeedback;
import software.amazon.awssdk.services.kendra.model.RelevanceType;
import software.amazon.awssdk.services.kendra.model.SubmitFeedbackRequest;
```

```
import software.amazon.awssdk.services.kendra.model.SubmitFeedbackResponse;

public class SubmitFeedbackExample {
    public static void main(String[] args) {
        KendraClient kendra = KendraClient.builder().build();

        SubmitFeedbackRequest submitFeedbackRequest = SubmitFeedbackRequest
            .builder()
            .indexId("IndexId")
            .queryId("QueryId")
            .clickFeedbackItems(
                ClickFeedback
                .builder()
                .clickTime(Instant.now())
                .resultId("ResultId")
                .build())
            .relevanceFeedbackItems(
                RelevanceFeedback
                .builder()
                .relevanceValue(RelevanceType.RELEVANT)
                .resultId("ResultId")
                .build())
            .build();

        SubmitFeedbackResponse response =
 kendra.submitFeedback(submitFeedbackRequest);

        System.out.println("Feedback is submitted");
    }
}
```

2. Exécutez le code. Une fois le commentaire envoyé, le code affiche un message.

Ajouter des synonymes personnalisés à un index

Pour ajouter des synonymes personnalisés à un index, vous devez les spécifier dans un fichier de thésaurus. Vous pouvez inclure des termes spécifiques à l'entreprise ou spécialisés lorsque vous Amazon Kendra utilisez des synonymes. Les synonymes anglais génériques, tels que`leader, head`, sont intégrés Amazon Kendra et ne doivent pas être inclus dans un fichier de thésaurus, y compris les synonymes génériques qui utilisent des traits d'union. Amazon Kendra prend en charge les synonymes pour tous les types de réponses, y compris les types de `DOCUMENT ANSWER` réponse `QUESTION_ANSWER` et/ou les types de réponse. Amazon Kendra ne prend actuellement pas en charge l'ajout de synonymes marqués comme des mots d'arrêt. Cela sera inclus dans une future version.

Amazon Kendra établit des corrélations entre les synonymes. Par exemple, l'utilisation de la paire `Dynamo, Amazon DynamoDB` de synonymes met Dynamo en Amazon Kendra corrélation avec. Amazon DynamoDB La requête « Qu'est-ce que la dynamo ? » renvoie ensuite un document tel que « Qu'est-ce que c'est Amazon DynamoDB ? ». Avec les synonymes, Amazon Kendra vous pouvez plus facilement détecter la corrélation.

Le fichier de thésaurus est un fichier texte stocké dans un Amazon S3 bucket. Consultez Ajouter un thésaurus à un index.

Le fichier de thésaurus utilise le format synonyme Solr. Amazon Kendra impose une limite au nombre de thésaurus par index. Consultez Quotas.

Les synonymes peuvent être utiles dans les scénarios suivants :

- Termes spécialisés qui ne sont pas des synonymes traditionnels en anglais, tels que`NLP, Natural Language Processing`.
- Noms propres avec des associations sémantiques complexes. Ce sont des noms que le grand public est peu susceptible de comprendre, par exemple, dans le domaine de l'apprentissage automatique. `cost, loss, model performance`
- Différentes formes de noms de produits, par exemple,`Elastic Compute Cloud, EC2`.
- Termes spécifiques au domaine ou à l'entreprise, tels que les noms de produits. Par exemple, `Route53, DNS`.

N'utilisez pas de synonymes dans les scénarios suivants :

- Synonymes génériques en anglais tels que`leader, head`. Ces synonymes ne sont pas spécifiques à un domaine, et l'utilisation de synonymes dans ces scénarios peut avoir des effets inattendus.
- Des erreurs typographiques telles que`teh => the`.
- Variantes morphologiques telles que le pluriel et le possessif des noms, la forme comparative et superlative des adjectifs, le passé, le participe passé et la forme progressive des verbes. Un exemple d'adjectifs comparatifs et superlatifs est. `good, better, best`
- Unigram (mot unique) arrête les mots tels que`WHO`. Les mots d'arrêt Unigram ne sont pas autorisés dans le thésaurus et sont exclus de la recherche. Par exemple, `WHO => World Health Organization` est rejeté. Vous pouvez `W.H.O.` toutefois l'utiliser comme synonyme, et vous pouvez utiliser des mots interrompus dans le cadre d'un synonyme comportant plusieurs mots. Par exemple, ce n'`of`est pas autorisé mais `United States of America` est accepté.

Les synonymes personnalisés facilitent la compréhension Amazon Kendra de la terminologie propre à votre entreprise en élargissant vos requêtes pour couvrir les synonymes spécifiques à votre entreprise. Bien que les synonymes puissent améliorer la précision de la recherche, il est important de comprendre comment les synonymes affectent la latence afin de pouvoir optimiser cela.

La règle générale en matière de synonymes est la suivante : plus le nombre de termes de votre requête mis en correspondance et développés avec des synonymes est élevé, plus l'impact potentiel sur le temps de latence est important. Parmi les autres facteurs qui influent sur la latence, citons la taille moyenne des documents indexés, la taille de votre index, les éventuels filtres appliqués aux résultats de recherche et la charge globale de votre Amazon Kendra index. Les requêtes qui ne correspondent à aucun synonyme ne sont pas affectées.

Une directive générale sur la façon dont les synonymes affectent le temps de latence :

Cas d'utilisation	Augmentation de la latence*
Requêtes classiques en langage naturel ou par mots clés de 3 à 5 mots chacune	Moins de 15 %
1 terme de requête est étendu à 3 synonymes	
Index d'environ 500 000 documents (en moyenne 10,48 Ko de texte extrait par	

Cas d'utilisation	Augmentation de la latence*
document) ou 30 000 paires de FAQ et de questions	

* Les performances varient en fonction de votre utilisation spécifique des synonymes et des configurations dans votre index. Il est préférable de tester les performances de recherche afin d'obtenir des points de référence plus précis pour votre cas d'utilisation spécifique.

Si votre thésaurus est volumineux, présente un taux d'extension à long terme élevé et que l'augmentation de la latence ne se situe pas dans les limites acceptables, vous pouvez essayer l'une des solutions suivantes ou les deux :

- Réduisez votre thésaurus pour réduire le taux d'extension (nombre de synonymes par terme).
- Réduisez la couverture globale des termes (nombre de lignes dans votre thésaurus).

Vous pouvez également augmenter la capacité de provisionnement (unités de stockage virtuelles) pour compenser l'augmentation de la latence.

Rubriques

- Création d'un fichier de thésaurus
- Ajouter un thésaurus à un index
- Mettre à jour un thésaurus
- Supprimer un thésaurus
- Points saillants dans les résultats de recherche

Création d'un fichier de thésaurus

Un fichier de Amazon Kendra thésaurus est un fichier codé en UTF-8 contenant une liste de synonymes au format de liste de synonymes Solr. La taille du fichier du thésaurus doit être inférieure à 5 Mo.

Il existe deux manières de spécifier les mappages de synonymes :

- Les synonymes bidirectionnels sont spécifiés sous forme de liste de termes séparés par des virgules. Si votre utilisateur interroge l'un des termes, tous les termes de la liste sont utilisés pour rechercher des documents, y compris le terme demandé d'origine.
- Les synonymes unidirectionnels sont spécifiés sous forme de termes séparés par le symbole « **=>** » entre eux pour associer les termes à leurs synonymes. Si votre utilisateur recherche un terme situé à gauche du symbole « **=>** », il est associé à un terme situé à droite pour rechercher des documents utilisant le synonyme. Il n'est pas mappé vice versa, ce qui le rend unidirectionnel.

Les synonymes eux-mêmes font la distinction majuscules/majuscules, mais les termes auxquels ils correspondent ne font pas la distinction majuscules/minuscules. Par exemple, `ML => Machine Learning` cela signifie que si votre utilisateur demande « ML » ou « ml » ou utilise un autre cas, il sera mappé sur « Machine Learning ». Si vous deviez mapper cela inversement`Machine Learning => ML`, alors « Machine Learning » ou « machine learning » ou un autre cas correspondrait à « ML ».

Un synonyme ne recherche pas de correspondance exacte entre les caractères spéciaux. Par exemple, si vous recherchez dead-letter-queue « », vous Amazon Kendra pouvez renvoyer des documents correspondant à la « file d'attente de lettres mortes » (sans tiret). Si vos documents contiennent des traits d'union, tels que dead-letter-queue « », Amazon Kendra traite les documents pendant la recherche pour supprimer les traits d'union. Pour les termes synonymes génériques en anglais qui sont intégrés à un fichier de thésaurus Amazon Kendra et ne devraient pas y être inclus, Amazon Kendra vous pouvez rechercher à la fois la version du terme avec tiret et la version sans trait d'union du terme. Par exemple, si vous recherchez « tiers » et « tiers », Amazon Kendra les documents qui correspondent à l'une ou l'autre version de ces termes sont renvoyés.

Pour les synonymes contenant des mots clés ou des mots couramment utilisés, Amazon Kendra renvoie des documents qui correspondent à des termes, y compris des mots d'arrêt. Par exemple, vous pouvez créer une règle de synonymes pour associer les termes « onboarding » et « onboarding ». Vous ne pouvez pas utiliser uniquement des mots d'arrêt pour les synonymes. Par exemple, si vous recherchez « on », vous Amazon Kendra ne pouvez pas renvoyer tous les documents contenant « on ».

Certaines règles relatives aux synonymes sont ignorées. Par exemple, `a => b` est une règle, mais elle `a => a` est ignorée et ne compte pas comme règle.

Le nombre de termes est le nombre de termes uniques contenus dans le fichier théaurus. Le fichier d'exemple ci-dessous inclut les termes`AWS CodeStar`,`ML`,`Machine Learning`,`autoscaling group`,`ASG`, et plus encore.

Il existe un nombre maximum de règles de synonymes par thésaurus et un maximum de synonymes par terme. Pour de plus amples informations, veuillez consulter Quotas pour Amazon Kendra.

L'exemple suivant montre un fichier de thésaurus avec des règles relatives aux synonymes. Chaque ligne contient une seule règle de synonymes. Les lignes vides et les commentaires sont ignorés.

```
# Lines starting with pound are comments and blank lines are ignored.

# Synonym relationships can be defined as unidirectional or bidirectional
 relationships.

# Unidirection relationships are represented by any term sequence
# on the left hand side (LHS) of "=>" followed by synonyms on the right hand side (RHS)
CodeStar => AWS CodeStar
# This will map CodeStar to AWS CodeStar, but not vice-versa

# To map terms vice versa
ML => Machine Learning
Machine Learning => ML

# Multiple synonym relationships may be defined in one line as well by comma
 seperation.
autoscaling group, ASG => Auto Scaling group, autoscaling
# The above is equivalent to:
# autoscaling group => Auto Scaling group, autoscaling
# ASG => Auto Scaling group, autoscaling

# Bi-directional synonyms are comma separated terms with no "=>"
DNS, Route53, Route 53
# DNS, Route53, and Route 53 map to one another and are interchangeable at match time
# The above is equivalent to:
# DNS => Route53, Route 53
# Route53 => DNS, Route 53
# Route 53 => DNS, Route53

# Overlapping LHS terms will be merged
Beta => Alpha
Beta => Gamma
Beta, Delta
# is equivalent to:
# Beta => Alpha, Gamma, Delta
# Delta => Beta
```

```
# Each line contains a single synonym rule.
# Synonym rule count is the total number of lines defining synonym relationships
# Term count is the total number of unique terms for all rules.
# Comments and blanks lines do not count.
```

Ajouter un thésaurus à un index

Les procédures suivantes indiquent comment ajouter un fichier de thésaurus contenant des synonymes à un index. Les effets de la mise à jour de votre fichier de thésaurus peuvent prendre jusqu'à 30 minutes. Pour plus d'informations sur le fichier de thésaurus, consultezCréation d'un fichier de thésaurus.

Console

Pour ajouter un thésaurus

1. Dans le volet de navigation de gauche, sous l'index dans lequel vous souhaitez ajouter une liste de synonymes, dans votre thésaurus, choisissez Synonymes.
2. Sur la page des synonymes, choisissez Ajouter un thésaurus.
3. Dans Définir le thésaurus, donnez un nom à votre thésaurus et une description facultative.
4. Dans les paramètres du thésaurus, indiquez le Amazon S3 chemin d'accès à votre fichier de thésaurus. La taille du fichier doit être inférieure à 5 Mo.
5. Pour le rôle IAM, sélectionnez un rôle ou sélectionnez Créer un nouveau rôle et spécifiez un nom de rôle pour créer un nouveau rôle. Amazon Kendra utilise ce rôle pour accéder à la Amazon S3 ressource en votre nom. Le rôle IAM porte le préfixe « AmazonKendra - ».
6. Choisissez Enregistrer pour enregistrer la configuration et ajouter le thésaurus. Une fois le thésaurus ingéré, il est actif et les synonymes sont mis en évidence dans les résultats. Les effets de votre fichier de thésaurus peuvent prendre jusqu'à 30 minutes.

CLI

Pour ajouter un thésaurus à un index avec le AWS CLI, appelez `create-thesaurus` :

```
aws kendra create-thesaurus \
--index-id index-id \
--name "thesaurus-name" \
--description "thesaurus-description" \
--source-s3-path "Bucket=bucket-name,Key=thesaurus/synonyms.txt" \
```

```
--role-arn role-arn
```

Appelez `list-thesauri` pour consulter la liste des thésaurus :

```
aws kendra list-thesauri \
--index-id index-id
```

Pour consulter les détails d'un thésaurus, appelez `describe-thesaurus` :

```
aws kendra describe-thesaurus \
--index-id index-id \
--index-id thesaurus-id
```

Les effets de votre fichier de thésaurus peuvent prendre jusqu'à 30 minutes.

Python

```
import boto3
from botocore.exceptions import ClientError
import pprint
import time

kendra = boto3.client("kendra")

print("Create a thesaurus")

thesaurus_name = "thesaurus-name"
thesaurus_description = "thesaurus-description"
thesaurus_role_arn = "role-arn"

index_id = "index-id"

s3_bucket_name = "bucket-name"
s3_key = "thesaurus-file"
source_s3_path= {
    'Bucket': s3_bucket_name,
    'Key': s3_key
}

try:
    thesaurus_response = kendra.create_thesaurus(
        Description = thesaurus_description,
        Name = thesaurus_name,
```

```
        RoleArn = thesaurus_role_arn,
        IndexId = index_id,
        SourceS3Path = source_s3_path
    )

    pprint.pprint(thesaurus_response)

    thesaurus_id = thesaurus_response["Id"]

    print("Wait for Kendra to create the thesaurus.")

    while True:
        # Get thesaurus description
        thesaurus_description = kendra.describe_thesaurus(
            Id = thesaurus_id,
            IndexId = index_id
        )
        # If status is not CREATING quit
        status = thesaurus_description["Status"]
        print("Creating thesaurus. Status: " + status)
        if status != "CREATING":
            break
        time.sleep(60)

except ClientError as e:
        print("%s" % e)

print("Program ends.")
```

Java

```
package com.amazonaws.kendra;

import software.amazon.awssdk.services.kendra.KendraClient;
import software.amazon.awssdk.services.kendra.model.CreateThesaurusRequest;
import software.amazon.awssdk.services.kendra.model.CreateThesaurusResponse;
import software.amazon.awssdk.services.kendra.model.DescribeThesaurusRequest;
import software.amazon.awssdk.services.kendra.model.DescribeThesaurusResponse;
import software.amazon.awssdk.services.kendra.model.S3Path;
import software.amazon.awssdk.services.kendra.model.ThesaurusStatus;

public class CreateThesaurusExample {
```

```
 public static void main(String[] args) throws InterruptedException {

   KendraClient kendra = KendraClient.builder().build();

   String thesaurusName = "thesaurus-name";
   String thesaurusDescription = "thesaurus-description";
   String thesaurusRoleArn = "role-arn";

   String s3BucketName = "bucket-name";
   String s3Key = "thesaurus-file";
   String indexId = "index-id";

   System.out.println(String.format("Creating a thesaurus named %s",
thesaurusName));
   CreateThesaurusRequest createThesaurusRequest = CreateThesaurusRequest
       .builder()
       .name(thesaurusName)
       .indexId(indexId)
       .description(thesaurusDescription)
       .roleArn(thesaurusRoleArn)
       .sourceS3Path(S3Path.builder()
           .bucket(s3BucketName)
           .key(s3Key)
           .build())
       .build();
   CreateThesaurusResponse createThesaurusResponse =
kendra.createThesaurus(createThesaurusRequest);
   System.out.println(String.format("Thesaurus response %s",
createThesaurusResponse));

   String thesaurusId = createThesaurusResponse.id();

   System.out.println(String.format("Waiting until the thesaurus with ID %s is
created.", thesaurusId));

   while (true) {
     DescribeThesaurusRequest describeThesaurusRequest =
DescribeThesaurusRequest.builder()
         .id(thesaurusId)
         .indexId(indexId)
         .build();
     DescribeThesaurusResponse describeThesaurusResponse =
kendra.describeThesaurus(describeThesaurusRequest);
     ThesaurusStatus status = describeThesaurusResponse.status();
```

```
      if (status != ThesaurusStatus.CREATING) {
        break;
      }

      TimeUnit.SECONDS.sleep(60);
    }

    System.out.println("Thesaurus creation is complete.");
  }
}
```

Mettre à jour un thésaurus

Vous pouvez modifier la configuration d'un thésaurus après sa création. Vous pouvez modifier des informations telles que le nom du thésaurus et les informations IAM. Vous pouvez également modifier l'emplacement du chemin Amazon S3 du fichier de thésaurus. Si vous modifiez le chemin du fichier de thésaurus, Amazon Kendra remplace le thésaurus existant par le thésaurus spécifié dans le chemin mis à jour.

Les effets de la mise à jour de votre fichier de thésaurus peuvent prendre jusqu'à 30 minutes.

> **Note**
>
> En cas d'erreurs de validation ou de syntaxe dans le fichier de thésaurus, le fichier de thésaurus précédemment téléchargé est conservé.

Les procédures suivantes indiquent comment modifier les détails du thésaurus.

Console

Pour modifier les détails du thésaurus

1. Dans le volet de navigation de gauche, sous l'index que vous souhaitez modifier, sélectionnez Synonymes.
2. Sur la page des synonymes, sélectionnez le thésaurus que vous souhaitez modifier, puis choisissez Modifier.
3. Sur la page Mettre à jour le thésaurus, mettez à jour les détails du thésaurus.

4. (Facultatif) Choisissez Modifier le chemin du fichier de thésaurus, puis spécifiez un Amazon S3 chemin d'accès au nouveau fichier de thésaurus. Votre fichier de thésaurus existant est remplacé par le fichier que vous spécifiez. Si vous ne modifiez pas le chemin, Amazon Kendra recharge le thésaurus à partir du chemin existant.

   Si vous sélectionnez Conserver le fichier de thésaurus actuel, Amazon Kendra cela ne recharge pas le fichier de thésaurus.
5. Choisissez Enregistrer pour enregistrer la configuration.

Vous pouvez également recharger le thésaurus à partir du chemin du thésaurus existant.

Pour recharger un thésaurus à partir d'un chemin existant

1. Dans le volet de navigation de gauche, sous l'index que vous souhaitez modifier, sélectionnez Synonymes.
2. Sur la page des synonymes, sélectionnez le thésaurus que vous souhaitez recharger, puis choisissez Actualiser.
3. Sur la page Recharger le fichier de thésaurus, confirmez que vous souhaitez actualiser le fichier de thésaurus.

CLI

Pour mettre à jour un thésaurus, appelez `update-thesaurus` :

```
aws kendra update-thesaurus \
--index-id index-id \
--name "thesaurus-name" \
--description "thesaurus-description" \
--source-s3-path "Bucket=bucket-name,Key=thesaurus/synonyms.txt" \
--role-arn role-arn
```

Python

```
import boto3
from botocore.exceptions import ClientError
import pprint
import time

kendra = boto3.client("kendra")
```

```
print("Update a thesaurus")

thesaurus_name = "thesaurus-name"
thesaurus_description = "thesaurus-description"
thesaurus_role_arn = "role-arn"

thesaurus_id = "thesaurus-id"
index_id = "index-id"

s3_bucket_name = "bucket-name"
s3_key = "thesaurus-file"
source_s3_path= {
    'Bucket': s3_bucket_name,
    'Key': s3_key
}

try:
    kendra.update_thesaurus(
        Id = thesaurus_id,
        IndexId = index_id,
        Description = thesaurus_description,
        Name = thesaurus_name,
        RoleArn = thesaurus_role_arn,
        SourceS3Path = source_s3_path
    )

    print("Wait for Kendra to update the thesaurus.")

    while True:
        # Get thesaurus description
        thesaurus_description = kendra.describe_thesaurus(
            Id = thesaurus_id,
            IndexId = index_id
        )
        # If status is not UPDATING quit
        status = thesaurus_description["Status"]
        print("Updating thesaurus. Status: " + status)
        if status != "UPDATING":
            break
        time.sleep(60)

except ClientError as e:
        print("%s" % e)
```

```
print("Program ends.")
```

Java

```
package com.amazonaws.kendra;

import software.amazon.awssdk.services.kendra.KendraClient;
import software.amazon.awssdk.services.kendra.model.UpdateThesaurusRequest;
import software.amazon.awssdk.services.kendra.model.DescribeThesaurusRequest;
import software.amazon.awssdk.services.kendra.model.DescribeThesaurusResponse;
import software.amazon.awssdk.services.kendra.model.S3Path;
import software.amazon.awssdk.services.kendra.model.ThesaurusStatus;

public class UpdateThesaurusExample {

  public static void main(String[] args) throws InterruptedException {

    KendraClient kendra = KendraClient.builder().build();

    String thesaurusName = "thesaurus-name";
    String thesaurusDescription = "thesaurus-description";
    String thesaurusRoleArn = "role-arn";

    String s3BucketName = "bucket-name";
    String s3Key = "thesaurus-file";

    String thesaurusId = "thesaurus-id";
    String indexId = "index-id";

    UpdateThesaurusRequest updateThesaurusRequest = UpdateThesaurusRequest
        .builder()
        .id(thesaurusId)
        .indexId(indexId)
        .name(thesaurusName)
        .description(thesaurusDescription)
        .roleArn(thesaurusRoleArn)
        .sourceS3Path(S3Path.builder()
            .bucket(s3BucketName)
            .key(s3Key)
            .build())
        .build();
    kendra.updateThesaurus(updateThesaurusRequest);
```

```
    System.out.println(String.format("Waiting until the thesaurus with ID %s is
 updated.", thesaurusId));

    // a new source s3 path requires re-consumption by Kendra
    // and so can take as long as a Create Thesaurus operation
    while (true) {
      DescribeThesaurusRequest describeThesaurusRequest =
 DescribeThesaurusRequest.builder()
          .id(thesaurusId)
          .indexId(indexId)
          .build();
      DescribeThesaurusResponse describeThesaurusResponse =
 kendra.describeThesaurus(describeThesaurusRequest);
      ThesaurusStatus status = describeThesaurusResponse.status();
      if (status != ThesaurusStatus.UPDATING) {
        break;
      }

      TimeUnit.SECONDS.sleep(60);
    }

    System.out.println("Thesaurus update is complete.");
  }
}
```

Supprimer un thésaurus

Les procédures suivantes indiquent comment supprimer un thésaurus.

Console

1. Dans le volet de navigation de gauche, sous l'index que vous souhaitez modifier, sélectionnez Synonymes.
2. Sur la page des synonymes, sélectionnez le thésaurus que vous souhaitez supprimer.
3. Sur la page détaillée du thésaurus, choisissez Supprimer, puis confirmez la suppression.

CLI

Pour supprimer un thésaurus dans un index avec le AWS CLI, appelez `delete-thesaurus` :

```
aws kendra delete-thesaurus \
--index-id index-id \
--id thesaurus-id
```

Python

```
import boto3
from botocore.exceptions import ClientError

kendra = boto3.client("kendra")

print("Delete a thesaurus")

thesaurus_id = "thesaurus-id"
index_id = "index-id"

try:
    kendra.delete_thesaurus(
        Id = thesaurus_id,
        IndexId = index_id
    )

except ClientError as e:
        print("%s" % e)

print("Program ends.")
```

Java

```
package com.amazonaws.kendra;

import software.amazon.awssdk.services.kendra.KendraClient;
import software.amazon.awssdk.services.kendra.model.DeleteThesaurusRequest;

public class DeleteThesaurusExample {

  public static void main(String[] args) throws InterruptedException {

    KendraClient kendra = KendraClient.builder().build();

    String thesaurusId = "thesaurus-id";
    String indexId = "index-id";
```

```
    DeleteThesaurusRequest updateThesaurusRequest = DeleteThesaurusRequest
        .builder()
        .id(thesaurusId)
        .indexId(indexId)
        .build();
    kendra.deleteThesaurus(updateThesaurusRequest);
  }
}
```

Points saillants dans les résultats de recherche

Le surlignage des synonymes est activé par défaut. Les informations de surlignage sont incluses dans les résultats des requêtes du Amazon Kendra SDK et de la CLI. Si vous interagissez à Amazon Kendra l'aide du SDK ou de la CLI, vous déterminez comment afficher les résultats.

Les surlignages synonymes auront le type `THESAURUS_SYNONYM` de surlignage. Pour plus d'informations sur les surlignages, consultez l'objet Highlight.

Tutoriel : Création d'une solution de recherche intelligente enrichie de métadonnées avec Amazon Kendra

Ce didacticiel vous explique comment créer une solution de recherche intelligente enrichie de métadonnées, basée sur le langage naturel et basée sur le langage naturel pour les données de votre entreprise à l'aide d'Amazon Kendra, Amazon Comprehend, Amazon Simple Storage Service (S3) et. AWS CloudShell

Amazon Kendra est un service de recherche intelligent capable de créer un index de recherche pour vos référentiels de données non structurés en langage naturel. Pour permettre à vos clients de trouver et de filtrer plus facilement les réponses pertinentes, vous pouvez utiliser Amazon Comprehend pour extraire les métadonnées de vos données et les intégrer dans votre index de recherche Amazon Kendra.

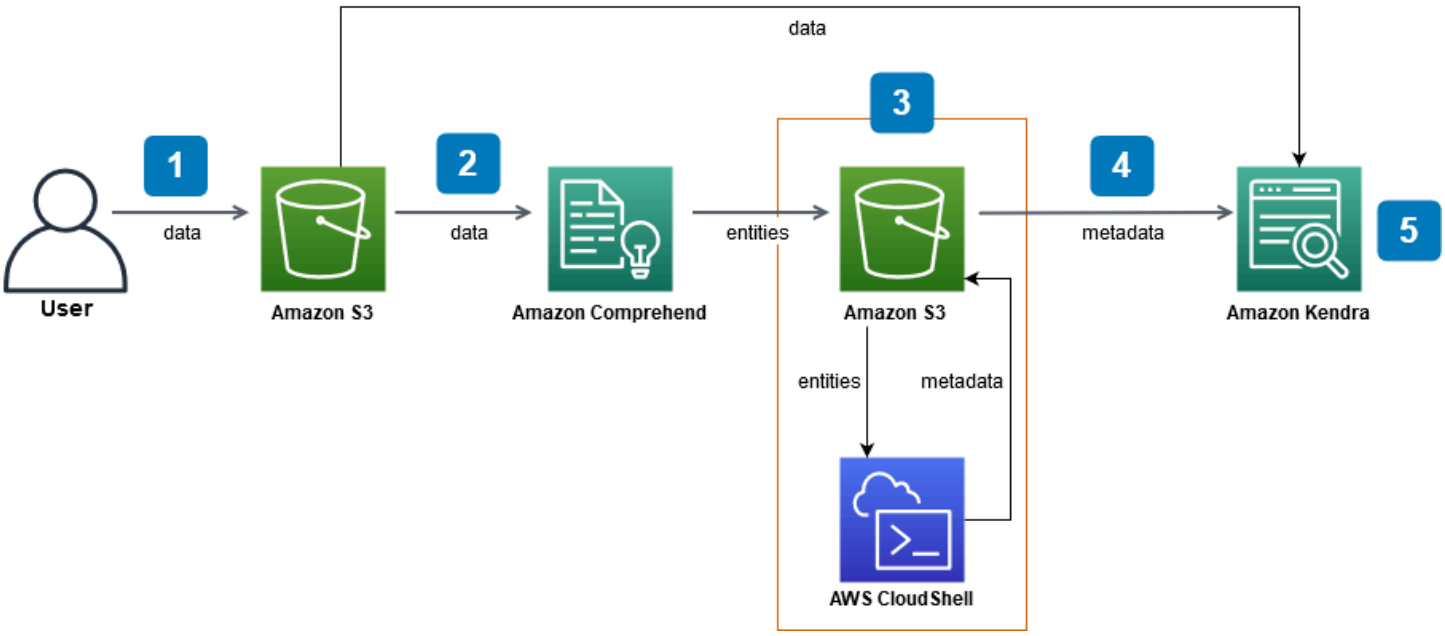
Amazon Comprehend est un service de traitement du langage naturel (NLP) capable d'identifier des entités. Les entités sont des références à des personnes, à des lieux, à des lieux, à des organisations et à des objets contenus dans vos données.

Ce didacticiel utilise un ensemble de données d'articles de presse pour extraire des entités, les convertir en métadonnées et les intégrer dans votre index Amazon Kendra pour y effectuer des recherches. Les métadonnées ajoutées vous permettent de filtrer les résultats de recherche à l'aide de n'importe quel sous-ensemble de ces entités et d'améliorer la précision de la recherche. En suivant ce didacticiel, vous apprendrez à créer une solution de recherche pour les données de votre entreprise sans aucune connaissance spécialisée en machine learning.

Ce didacticiel explique comment créer votre solution de recherche en suivant les étapes suivantes :

1. Stockage d'un ensemble de données d'articles de presse dans Amazon S3.
2. Utiliser Amazon Comprehend pour extraire des entités de vos données.
3. Exécution d'un script Python 3 pour convertir les entités au format de métadonnées d'index Amazon Kendra et stockage de ces métadonnées dans S3.
4. Création d'un index de recherche Amazon Kendra et ingestion des données et des métadonnées.
5. Interrogation de l'index de recherche.

Le schéma suivant montre le flux de travail :



Temps estimé pour terminer ce didacticiel : 1 heure

Coût estimé : Certaines des actions décrites dans ce didacticiel entraînent des frais sur votre AWS compte. Pour plus d'informations sur le coût de chaque service, consultez les pages de prix d'Amazon S3, Amazon Comprehend et Amazon AWS CloudShellKendra.

Rubriques

- Prérequis
- Étape 1 : ajout de documents à Amazon S3
- Étape 2 : Exécution d'une tâche d'analyse d'entités sur Amazon Comprehend
- Étape 3 : Formatage du résultat de l'analyse des entités sous forme de métadonnées Amazon Kendra
- Étape 4 : Création d'un index Amazon Kendra et ingestion des métadonnées
- Étape 5 : Interrogation de l'index Amazon Kendra
- Étape 6 : Nettoyage

Prérequis

Pour suivre ce didacticiel, vous avez besoin des ressources suivantes :

- Un AWS compte. Si vous n'avez pas de AWS compte, suivez les étapes décrites dans Configuration d'Amazon Kendra pour configurer votre AWS compte.

- Un ordinateur de développement exécutant Windows, macOS ou Linux pour accéder à la console AWS de gestion. Pour plus d'informations, consultez la section Configuration de la console AWS de gestion.
- Un utilisateur AWS Identity and Access Management(IAM). Pour savoir comment configurer un utilisateur et un groupe IAM pour votre compte, consultez la section Getting Started du guide de l'utilisateur IAM.

  Si vous utilisez le AWS Command Line Interface, vous devez également associer la politique suivante à votre utilisateur IAM afin de lui accorder les autorisations de base requises pour suivre ce didacticiel.

  Pour plus d'informations, consultez les sections Création de politiques IAM et Ajout et suppression d'autorisations d'identité IAM.
- La liste des services AWS régionaux. Pour réduire le temps de latence, vous devez choisir la AWS région la plus proche de votre situation géographique qui est prise en charge à la fois par Amazon Comprehend et Amazon Kendra.
- (Facultatif) Un AWS Key Management Service. Bien que ce didacticiel n'utilise pas le chiffrement, vous souhaiterez peut-être utiliser les meilleures pratiques de chiffrement pour votre cas d'utilisation spécifique.
- (Facultatif) Un Amazon Virtual Private Cloud. Bien que ce didacticiel n'utilise pas de VPC, vous souhaiterez peut-être utiliser les meilleures pratiques en matière de VPC pour garantir la sécurité des données pour votre cas d'utilisation spécifique.

Étape 1 : ajout de documents à Amazon S3

Avant d'exécuter une tâche d'analyse des entités Amazon Comprehend sur votre ensemble de données, vous devez créer un compartiment Amazon S3 pour héberger les données, les métadonnées et les résultats de l'analyse des entités Amazon Comprehend.

Rubriques

- Téléchargement de l'exemple de jeu de données
- Création d'un compartiment Amazon S3
- Création de dossiers de données et de métadonnées dans votre compartiment S3
- Téléchargement des données d'entrée

Téléchargement de l'exemple de jeu de données

Avant qu'Amazon Comprehend puisse exécuter une tâche d'analyse d'entités sur vos données, vous devez télécharger et extraire le jeu de données, puis le charger dans un compartiment S3.

Pour télécharger et extraire le jeu de données (console)

1. Téléchargez le dossier tutorial-dataset.zip sur votre appareil.
2. Extrayez le `tutorial-dataset` dossier pour y accéder. `data`

Pour télécharger et extraire le jeu de données (Terminal)

1. Pour le téléchargertutorial-dataset, exécutez la commande suivante dans une fenêtre de terminal :

   Linux

   ```
   curl -o path/tutorial-dataset.zip https://docs.aws.amazon.com/kendra/latest/dg/
   samples/tutorial-dataset.zip
   ```

   Où :

   - *path/*est le chemin du fichier local vers l'emplacement dans lequel vous souhaitez enregistrer le dossier zip.

   macOS

   ```
   curl -o path/tutorial-dataset.zip https://docs.aws.amazon.com/kendra/latest/dg/
   samples/tutorial-dataset.zip
   ```

   Où :

   - *path/*est le chemin du fichier local vers l'emplacement dans lequel vous souhaitez enregistrer le dossier zip.

Windows

```
curl -o path/tutorial-dataset.zip https://docs.aws.amazon.com/kendra/latest/dg/
samples/tutorial-dataset.zip
```

Où :

- *path/*est le chemin du fichier local vers l'emplacement dans lequel vous souhaitez enregistrer le dossier zip.

2. Pour extraire les données du dossier zip, exécutez la commande suivante dans la fenêtre du terminal :

Linux

```
unzip path/tutorial-dataset.zip -d path/
```

Où :

- *path/*est le chemin de fichier local vers le dossier zip que vous avez enregistré.

macOS

```
unzip path/tutorial-dataset.zip -d path/
```

Où :

- *path/*est le chemin de fichier local vers le dossier zip que vous avez enregistré.

Windows

```
tar -xf path/tutorial-dataset.zip -C path/
```

Où :

- *path/*est le chemin de fichier local vers le dossier zip que vous avez enregistré.

À la fin de cette étape, vous devriez avoir les fichiers extraits dans un dossier décompressé appelé`tutorial-dataset`. Ce dossier contient un `README` fichier avec une attribution open source Apache 2.0 et un dossier appelé `data` contenant le jeu de données pour ce didacticiel. Le jeu de données se compose de 100 fichiers avec des `.story` extensions.

Création d'un compartiment Amazon S3

Après avoir téléchargé et extrait le dossier de données d'exemple, vous le stockez dans un compartiment Amazon S3.

> **Important**
>
> Le nom d'un compartiment Amazon S3 doit être unique pour tous AWS.

Pour créer un compartiment S3 (console)

1. Connectez-vous à la console Amazon S3 AWS Management Console et ouvrez-la à l'adresse https://console.aws.amazon.com/s3/.
2. Dans Buckets, choisissez Create bucket.
3. Pour Nom de compartiment, entrez un nom unique.
4. Pour Région, choisissez la AWS région dans laquelle vous souhaitez créer le bucket.

   > **Note**
   >
   > Vous devez choisir une région qui prend en charge à la fois Amazon Comprehend et Amazon Kendra. Vous ne pouvez pas modifier la région d'un bucket après l'avoir créé.

5. Conservez les paramètres par défaut pour les paramètres de blocage de l'accès public pour ce compartiment, le contrôle de version du compartiment et les balises.
6. Pour le chiffrement par défaut, choisissez Désactiver.
7. Conservez les paramètres par défaut pour les paramètres avancés.
8. Vérifiez la configuration de votre compartiment, puis choisissez Create bucket.

Pour créer un compartiment S3 (AWS CLI)

1. Pour créer un compartiment S3, utilisez la commande create-bucket dans le : AWS CLI

Linux

```
aws s3api create-bucket \
        --bucket amzn-s3-demo-bucket \
        --region aws-region \
        --create-bucket-configuration LocationConstraint=aws-region
```

Où :

- amzn-s3-demo-bucket est le nom de votre bucket,
- *aws-region*est la région dans laquelle vous souhaitez créer votre bucket.

macOS

```
aws s3api create-bucket \
        --bucket amzn-s3-demo-bucket \
        --region aws-region \
        --create-bucket-configuration LocationConstraint=aws-region
```

Où :

- amzn-s3-demo-bucket est le nom de votre bucket,
- *aws-region*est la région dans laquelle vous souhaitez créer votre bucket.

Windows

```
aws s3api create-bucket ^
        --bucket amzn-s3-demo-bucket ^
        --region aws-region ^
        --create-bucket-configuration LocationConstraint=aws-region
```

Où :

- amzn-s3-demo-bucket est le nom de votre bucket,
- *aws-region*est la région dans laquelle vous souhaitez créer votre bucket.

> **Note**
>
> Vous devez choisir une région qui prend en charge à la fois Amazon Comprehend et Amazon Kendra. Vous ne pouvez pas modifier la région d'un bucket après l'avoir créé.

2. Pour vous assurer que votre bucket a été créé avec succès, utilisez la commande list :

   Linux

   ```
   aws s3 ls
   ```

   macOS

   ```
   aws s3 ls
   ```

   Windows

   ```
   aws s3 ls
   ```

Création de dossiers de données et de métadonnées dans votre compartiment S3

Après avoir créé votre compartiment S3, vous créez des dossiers de données et de métadonnées à l'intérieur de celui-ci.

Pour créer des dossiers dans votre compartiment S3 (console)

1. Ouvrez la console Amazon S3 à l'adresse https://console.aws.amazon.com/s3/.
2. Dans Buckets, cliquez sur le nom de votre bucket dans la liste des buckets.
3. Dans l'onglet Objets, choisissez Créer un dossier.
4. Pour le nouveau nom du dossier, entrez**`data`**.
5. Pour les paramètres de chiffrement, choisissez Désactiver.
6. Choisissez Créer un dossier.
7. Répétez les étapes 3 à 6 pour créer un autre dossier destiné à stocker les métadonnées Amazon Kendra et nommez le dossier créé à l'étape 4. **`metadata`**

Pour créer des dossiers dans votre compartiment S3 (AWS CLI)

1. Pour créer le `data` dossier dans votre compartiment S3, utilisez la commande put-object dans : AWS CLI

   Linux

   ```
   aws s3api put-object \
           --bucket amzn-s3-demo-bucket \
           --key data/
   ```

   Où :

   - amzn-s3-demo-bucket est le nom de votre bucket.

   macOS

   ```
   aws s3api put-object \
           --bucket amzn-s3-demo-bucket \
           --key data/
   ```

   Où :

   - amzn-s3-demo-bucket est le nom de votre bucket.

   Windows

   ```
   aws s3api put-object ^
           --bucket amzn-s3-demo-bucket ^
           --key data/
   ```

   Où :

   - amzn-s3-demo-bucket est le nom de votre bucket.

2. Pour créer le `metadata` dossier dans votre compartiment S3, utilisez la commande put-object dans : AWS CLI

Linux

```
aws s3api put-object \
        --bucket amzn-s3-demo-bucket \
        --key metadata/
```

Où :

- amzn-s3-demo-bucket est le nom de votre bucket.

macOS

```
aws s3api put-object \
        --bucket amzn-s3-demo-bucket \
        --key metadata/
```

Où :

- amzn-s3-demo-bucket est le nom de votre bucket.

Windows

```
aws s3api put-object ^
        --bucket amzn-s3-demo-bucket ^
        --key metadata/
```

Où :

- amzn-s3-demo-bucket est le nom de votre bucket.

3. Pour vous assurer que vos dossiers ont été créés correctement, vérifiez le contenu de votre bucket à l'aide de la commande list :

Linux

```
aws s3 ls s3://amzn-s3-demo-bucket/
```

Où :

- amzn-s3-demo-bucket est le nom de votre bucket.

macOS

```
aws s3 ls s3://amzn-s3-demo-bucket/
```

Où :

- amzn-s3-demo-bucket est le nom de votre bucket.

Windows

```
aws s3 ls s3://amzn-s3-demo-bucket/
```

Où :

- amzn-s3-demo-bucket est le nom de votre bucket.

Téléchargement des données d'entrée

Après avoir créé vos dossiers de données et de métadonnées, vous chargez l'exemple de jeu de données dans le `data` dossier.

Pour télécharger l'exemple de jeu de données dans le dossier de données (console)

1. Ouvrez la console Amazon S3 à l'adresse https://console.aws.amazon.com/s3/.
2. Dans Buckets, cliquez sur le nom de votre bucket dans la liste des buckets, puis cliquez sur. `data`
3. Choisissez Télécharger, puis Ajouter des fichiers.
4. Dans la boîte de dialogue, accédez au `data` dossier situé dans le `tutorial-dataset` dossier de votre appareil local, sélectionnez tous les fichiers, puis choisissez Ouvrir.
5. Conservez les paramètres par défaut pour la destination, les autorisations et les propriétés.
6. Choisissez Charger.

Pour télécharger l'exemple de jeu de données dans le dossier de données (AWS CLI)

1. Pour télécharger les exemples de données dans le `data` dossier, utilisez la commande de copie dans AWS CLI :

   Linux

   ```
   aws s3 cp path/tutorial-dataset/data s3://amzn-s3-demo-bucket/data/ --recursive
   ```

   Où :

   - *path/*est le chemin du fichier vers le `tutorial-dataset` dossier sur votre appareil,
   - amzn-s3-demo-bucket est le nom de votre bucket.

   macOS

   ```
   aws s3 cp path/tutorial-dataset/data s3://amzn-s3-demo-bucket/data/ --recursive
   ```

   Où :

   - *path/*est le chemin du fichier vers le `tutorial-dataset` dossier sur votre appareil,
   - amzn-s3-demo-bucket est le nom de votre bucket.

   Windows

   ```
   aws s3 cp path/tutorial-dataset/data s3://amzn-s3-demo-bucket/data/ --recursive
   ```

   Où :

   - *path/*est le chemin du fichier vers le `tutorial-dataset` dossier sur votre appareil,
   - amzn-s3-demo-bucket est le nom de votre bucket.

2. Pour vous assurer que les fichiers de votre ensemble de données ont été correctement chargés `data` dans votre dossier, utilisez la commande list dans le AWS CLI :

   Linux

   ```
   aws s3 ls s3://amzn-s3-demo-bucket/data/
   ```

Où :

- amzn-s3-demo-bucket est le nom de votre compartiment S3.

macOS

```
aws s3 ls s3://amzn-s3-demo-bucket/data/
```

Où :

- amzn-s3-demo-bucket est le nom de votre compartiment S3.

Windows

```
aws s3 ls s3://amzn-s3-demo-bucket/data/
```

Où :

- amzn-s3-demo-bucket est le nom de votre compartiment S3.

À la fin de cette étape, vous disposez d'un compartiment S3 dans lequel votre ensemble de données est stocké `data` dans le dossier, et d'un `metadata` dossier vide dans lequel seront stockées vos métadonnées Amazon Kendra.

Étape 2 : Exécution d'une tâche d'analyse d'entités sur Amazon Comprehend

Après avoir stocké l'exemple de jeu de données dans votre compartiment S3, vous exécutez une tâche d'analyse des entités Amazon Comprehend pour extraire les entités de vos documents. Ces entités formeront des attributs personnalisés d'Amazon Kendra et vous aideront à filtrer les résultats de recherche sur votre index. Pour plus d'informations, consultez la section Détecter les entités.

Rubriques

- Exécution d'une tâche d'analyse d'entités Amazon Comprehend

Exécution d'une tâche d'analyse d'entités Amazon Comprehend

Pour extraire des entités de votre ensemble de données, vous devez exécuter une tâche d'analyse d'entités Amazon Comprehend.

Si vous utilisez la AWS CLI au cours de cette étape, vous devez d'abord créer et associer un rôle et une politique AWS IAM pour Amazon Comprehend, puis exécuter une tâche d'analyse des entités. Pour exécuter une tâche d'analyse d'entités sur vos échantillons de données, Amazon Comprehend a besoin des éléments suivants :

- un rôle AWS Identity and Access Management (IAM) qui le reconnaît comme une entité de confiance
- une politique AWS IAM attachée au rôle IAM qui lui donne l'autorisation d'accéder à votre compartiment S3

Pour plus d'informations, consultez Comment Amazon Comprehend fonctionne avec les politiques IAM et basées sur l'identité pour Amazon Comprehend.

Pour exécuter une tâche d'analyse d'entités Amazon Comprehend (console)

1. Ouvrez la console Amazon Comprehend à l'adresse. https://console.aws.amazon.com/comprehend/

   > **Important**
   >
   > Assurez-vous que vous vous trouvez dans la même région que celle dans laquelle vous avez créé votre compartiment Amazon S3. Si vous vous trouvez dans une autre région, choisissez la AWS région dans laquelle vous avez créé votre compartiment S3 dans le sélecteur de région situé dans la barre de navigation supérieure.

2. Choisissez Launch Amazon Comprehend.
3. Dans le volet de navigation de gauche, sélectionnez Analysis jobs.
4. Choisissez Créer une tâche.
5. Dans la section Paramètres du job, procédez comme suit :

   a. Pour Nom, saisissez **`data-entities-analysis`**.

   b. Pour le type d'analyse, sélectionnez Entités.

c. Dans Langue, choisissez l'anglais.

d. Gardez le chiffrement Job désactivé.

6. Dans la section Données d'entrée, procédez comme suit :

a. Pour Source de données, sélectionnez Mes documents.

b. Pour l'emplacement S3, choisissez Browse S3.

c. Pour Choose resources, cliquez sur le nom de votre bucket dans la liste des buckets.

d. Pour Objets, sélectionnez le bouton d'option pour `data` et choisissez Choisir.

e. Pour Format d'entrée, choisissez Un document par fichier.

7. Dans la section Données de sortie, procédez comme suit :

a. Pour l'emplacement S3, choisissez Browse S3, puis sélectionnez la case d'option correspondant à votre bucket dans la liste des buckets et choisissez Choose.

b. Maintenez le chiffrement désactivé.

8. Dans la section Autorisations d'accès, procédez comme suit :

a. Pour le rôle IAM, choisissez Create an IAM role.

b. Pour les autorisations d'accès, choisissez les compartiments S3 d'entrée et de sortie.

c. Dans le champ Suffixe du nom, entrez**`comprehend-role`**. Ce rôle permet d'accéder à votre compartiment Amazon S3.

9. Conservez les paramètres VPC par défaut.

10. Choisissez Créer une tâche.

Pour exécuter une tâche d'analyse d'entités Amazon Comprehend ()AWS CLI

1. Pour créer et associer un rôle IAM à Amazon Comprehend qui le reconnaisse comme une entité de confiance, procédez comme suit :

a. Enregistrez la politique de confiance suivante sous forme de fichier JSON appelé `comprehend-trust-policy.json` dans un éditeur de texte sur votre appareil local.

JSON

```
{
  "Version": "2012-10-17",
```

```
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "Service": "comprehend.amazonaws.com"
      },
      "Action": "sts:AssumeRole"
    }
  ]
}
```

b. Pour créer un rôle IAM appelé `comprehend-role` et y joindre votre `comprehend-trust-policy.json` fichier enregistré, utilisez la commande create-role :

Linux

```
aws iam create-role \
          --role-name comprehend-role \
          --assume-role-policy-document file://path/comprehend-trust-
policy.json
```

Où :

- *path/*est le chemin du fichier vers votre `comprehend-trust-policy.json` appareil local.

macOS

```
aws iam create-role \
          --role-name comprehend-role \
          --assume-role-policy-document file://path/comprehend-trust-
policy.json
```

Où :

- *path/*est le chemin du fichier vers votre `comprehend-trust-policy.json` appareil local.

Windows

```
aws iam create-role ^
         --role-name comprehend-role ^
         --assume-role-policy-document file://path/comprehend-trust-
policy.json
```

Où :

- *path/*est le chemin du fichier vers votre `comprehend-trust-policy.json` appareil local.

c. Copiez l'Amazon Resource Name (ARN) dans votre éditeur de texte et enregistrez-le localement sous le nom de`comprehend-role-arn`.

> **Note**
>
> Le format de l'ARN est similaire à*arn:aws:iam::123456789012:role/comprehend-role*. Vous avez besoin de l'ARN sous lequel vous avez enregistré `comprehend-role-arn` pour exécuter la tâche d'analyse Amazon Comprehend.

2. Pour créer et associer une politique IAM à votre rôle IAM qui lui accorde l'autorisation d'accéder à votre compartiment S3, procédez comme suit :

a. Enregistrez la politique de confiance suivante sous forme de fichier JSON appelé `comprehend-S3-access-policy.json` dans un éditeur de texte sur votre appareil local.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Action": [
                "s3:GetObject"
            ],
            "Resource": [
                "arn:aws:s3:::amzn-s3-demo-bucket/*"
            ],
```

```
            "Effect": "Allow"
        },
        {
            "Action": [
                "s3:ListBucket"
            ],
            "Resource": [
                "arn:aws:s3:::amzn-s3-demo-bucket"
            ],
            "Effect": "Allow"
        },
        {
            "Action": [
                "s3:PutObject"
            ],
            "Resource": [
                "arn:aws:s3:::amzn-s3-demo-bucket/*"
            ],
            "Effect": "Allow"
        }
    ]
}
```

b. Pour créer une politique IAM appelée `comprehend-S3-access-policy` pour accéder à votre compartiment S3, utilisez la commande create-policy :

Linux

```
aws iam create-policy \
         --policy-name comprehend-S3-access-policy \
         --policy-document file://path/comprehend-S3-access-policy.json
```

Où :

- *path/*est le chemin du fichier vers votre `comprehend-S3-access-policy.json` appareil local.

macOS

```
aws iam create-policy \
         --policy-name comprehend-S3-access-policy \
```

```
        --policy-document file://path/comprehend-S3-access-policy.json
```

Où :

- *path/*est le chemin du fichier vers votre `comprehend-S3-access-policy.json` appareil local.

Windows

```
aws iam create-policy ^
        --policy-name comprehend-S3-access-policy ^
        --policy-document file://path/comprehend-S3-access-policy.json
```

Où :

- *path/*est le chemin du fichier vers votre `comprehend-S3-access-policy.json` appareil local.

c. Copiez l'Amazon Resource Name (ARN) dans votre éditeur de texte et enregistrez-le localement sous le nom de`comprehend-S3-access-arn`.

> **Note**
>
> Le format de l'ARN est similaire à*arn:aws:iam::123456789012:role/comprehend-S3-access-policy*. Vous avez besoin de l'ARN sous lequel vous avez enregistré `comprehend-S3-access-arn` pour l'associer `comprehend-S3-access-policy` à votre rôle IAM.

d. Pour l'associer `comprehend-S3-access-policy` à votre rôle IAM, utilisez la attach-role-policycommande :

Linux

```
aws iam attach-role-policy \
        --policy-arn policy-arn \
        --role-name comprehend-role
```

Où :

- *policy-arn*est l'ARN sous lequel vous avez enregistrécomprehend-S3-access-arn.

macOS

```
aws iam attach-role-policy \
         --policy-arn policy-arn \
         --role-name comprehend-role
```

Où :

- *policy-arn*est l'ARN sous lequel vous avez enregistrécomprehend-S3-access-arn.

Windows

```
aws iam attach-role-policy ^
         --policy-arn policy-arn ^
         --role-name comprehend-role
```

Où :

- *policy-arn*est l'ARN sous lequel vous avez enregistrécomprehend-S3-access-arn.

3. Pour exécuter une tâche d'analyse d'entités Amazon Comprehend, utilisez la start-entities-detection-jobcommande suivante :

Linux

```
aws comprehend start-entities-detection-job \
        --input-data-config S3Uri=s3://amzn-s3-demo-bucket/
data/,InputFormat=ONE_DOC_PER_FILE \
        --output-data-config S3Uri=s3://amzn-s3-demo-bucket/ \
        --data-access-role-arn role-arn \
        --job-name data-entities-analysis \
        --language-code en \
        --region aws-region
```

Où :

- amzn-s3-demo-bucket est le nom de votre compartiment S3,
- *role-arn*est l'ARN sous lequel vous avez enregistré`comprehend-role-arn`,
- *aws-region*est votre AWS région.

macOS

```
aws comprehend start-entities-detection-job \
        --input-data-config S3Uri=s3://amzn-s3-demo-bucket/
data/,InputFormat=ONE_DOC_PER_FILE \
        --output-data-config S3Uri=s3://amzn-s3-demo-bucket/ \
        --data-access-role-arn role-arn \
        --job-name data-entities-analysis \
        --language-code en \
        --region aws-region
```

Où :

- amzn-s3-demo-bucket est le nom de votre compartiment S3,
- *role-arn*est l'ARN sous lequel vous avez enregistré`comprehend-role-arn`,
- *aws-region*est votre AWS région.

Windows

```
aws comprehend start-entities-detection-job ^
        --input-data-config S3Uri=s3://amzn-s3-demo-bucket/
data/,InputFormat=ONE_DOC_PER_FILE ^
        --output-data-config S3Uri=s3://amzn-s3-demo-bucket/ ^
        --data-access-role-arn role-arn ^
        --job-name data-entities-analysis ^
        --language-code en ^
        --region aws-region
```

Où :

- amzn-s3-demo-bucket est le nom de votre compartiment S3,

- *role-arn*est l'ARN sous lequel vous avez enregistré`comprehend-role-arn`,
- *aws-region*est votre AWS région.

4. Copiez l'analyse des entités `JobId` et enregistrez-la dans un éditeur de texte sous`comprehend-job-id`. `JobId`Cela vous permet de suivre l'état de votre tâche d'analyse des entités.
5. Pour suivre la progression de votre tâche d'analyse des entités, utilisez la describe-entities-detection-jobcommande :

   Linux

   ```
   aws comprehend describe-entities-detection-job \
           --job-id entities-job-id \
           --region aws-region
   ```

   Où :

   - *entities-job-id*est votre sauvegarde`comprehend-job-id`,
   - *aws-region*est votre AWS région.

   macOS

   ```
   aws comprehend describe-entities-detection-job \
           --job-id entities-job-id \
           --region aws-region
   ```

   Où :

   - *entities-job-id*est votre sauvegarde`comprehend-job-id`,
   - *aws-region*est votre AWS région.

   Windows

   ```
   aws comprehend describe-entities-detection-job ^
           --job-id entities-job-id ^
           --region aws-region
   ```

   Où :

- *entities-job-id*est votre sauvegardecomprehend-job-id,
- *aws-region*est votre AWS région.

Cela peut prendre plusieurs minutes pour que le passe JobStatus àCOMPLETED.

À la fin de cette étape, Amazon Comprehend stocke les résultats de l'analyse des entités sous forme de output.tar.gz fichier compressé dans un output dossier généré automatiquement dans votre compartiment S3. Assurez-vous que le statut de votre tâche d'analyse est terminé avant de passer à l'étape suivante.

Étape 3 : Formatage du résultat de l'analyse des entités sous forme de métadonnées Amazon Kendra

Pour convertir les entités extraites par Amazon Comprehend au format de métadonnées requis par un index Amazon Kendra, vous devez exécuter un script Python 3. Les résultats de la conversion sont stockés dans le metadata dossier de votre compartiment Amazon S3.

Pour plus d'informations sur le format et la structure des métadonnées Amazon Kendra, consultez la section Métadonnées du document S3.

Rubriques

- Téléchargement et extraction de la sortie Amazon Comprehend
- Téléchargement de la sortie dans le compartiment S3
- Conversion de la sortie au format de métadonnées Amazon Kendra
- Nettoyage de votre compartiment Amazon S3

Téléchargement et extraction de la sortie Amazon Comprehend

Pour formater le résultat de l'analyse des entités Amazon Comprehend, vous devez d'abord télécharger l'archive d'analyse des entités Amazon Comprehend et extraire le fichier d'output.tar.gzanalyse des entités.

Pour télécharger et extraire le fichier de sortie (console)

1. Dans le volet de navigation de la console Amazon Comprehend, accédez à Analysis jobs.

2. Choisissez votre tâche d'analyse d'entitésdata-entities-analysis.
3. Sous Sortie, choisissez le lien affiché à côté de Emplacement des données de sortie. Cela vous redirige vers l'output.tar.gzarchive de votre compartiment S3.
4. Dans l'onglet Vue d'ensemble, choisissez Télécharger.

> **Tip**
>
> Les résultats de toutes les tâches d'analyse Amazon Comprehend portent le même nom. Le fait de renommer votre archive vous permettra de la suivre plus facilement.

5. Décompressez et extrayez le fichier Amazon Comprehend téléchargé sur votre appareil.

Pour télécharger et extraire le fichier de sortie (AWS CLI)

1. Pour accéder au nom du dossier généré automatiquement par Amazon Comprehend dans votre compartiment S3 qui contient les résultats de la tâche d'analyse des entités, utilisez la commande suivante : describe-entities-detection-job

   Linux

   ```
   aws comprehend describe-entities-detection-job \
             --job-id entities-job-id \
             --region aws-region
   ```

   Où :

   - *entities-job-id*est votre sauvegarde comprehend-job-id depuisthe section called “Étape 2 : Détection des entités”,
   - *aws-region*est votre AWS région.

   macOS

   ```
   aws comprehend describe-entities-detection-job \
             --job-id entities-job-id \
             --region aws-region
   ```

   Où :

- *entities-job-id*est votre sauvegarde `comprehend-job-id` depuisthe section called "Étape 2 : Détection des entités",
- *aws-region*est votre AWS région.

Windows

```
aws comprehend describe-entities-detection-job ^
          --job-id entities-job-id ^
          --region aws-region
```

Où :

- *entities-job-id*est votre sauvegarde `comprehend-job-id` depuisthe section called "Étape 2 : Détection des entités",
- *aws-region*est votre AWS région.

2. À partir de l'`OutputDataConfig`objet figurant dans la description de travail de votre entité, copiez et enregistrez la `S3Uri` valeur dans `comprehend-S3uri` un éditeur de texte.

> **Note**
>
> Le format de `S3Uri` la valeur est similaire à*s3://amzn-s3-demo-bucket/.../output/output.tar.gz*.

3. Pour télécharger l'archive de sortie des entités, utilisez la commande copy :

Linux

```
aws s3 cp s3://amzn-s3-demo-bucket/.../output/output.tar.gz path/output.tar.gz
```

Où :

- *s3://amzn-s3-demo-bucket/.../output/output.tar.gz*est la `S3Uri` valeur sous laquelle vous avez enregistré`comprehend-S3uri`,
- *path/*est le répertoire local dans lequel vous souhaitez enregistrer la sortie.

macOS

```
aws s3 cp s3://amzn-s3-demo-bucket/.../output/output.tar.gz path/output.tar.gz
```

Où :

- *s3://amzn-s3-demo-bucket/.../output/output.tar.gz*est la S3Uri valeur sous laquelle vous avez enregistrécomprehend-S3uri,
- *path/*est le répertoire local dans lequel vous souhaitez enregistrer la sortie.

Windows

```
aws s3 cp s3://amzn-s3-demo-bucket/.../output/output.tar.gz path/output.tar.gz
```

Où :

- *s3://amzn-s3-demo-bucket/.../output/output.tar.gz*est la S3Uri valeur sous laquelle vous avez enregistrécomprehend-S3uri,
- *path/*est le répertoire local dans lequel vous souhaitez enregistrer la sortie.

4. Pour extraire la sortie des entités, exécutez la commande suivante dans une fenêtre de terminal :

Linux

```
tar -xf path/output.tar.gz -C path/
```

Où :

- *path/*est le chemin d'accès à l'output.tar.gzarchive téléchargée sur votre appareil local.

macOS

```
tar -xf path/output.tar.gz -C path/
```

Où :

- *path/*est le chemin d'accès à l'`output.tar.gz`archive téléchargée sur votre appareil local.

Windows

```
tar -xf path/output.tar.gz -C path/
```

Où :

- *path/*est le chemin d'accès à l'`output.tar.gz`archive téléchargée sur votre appareil local.

À la fin de cette étape, vous devriez avoir un fichier sur votre appareil appelé `output` contenant une liste des entités identifiées par Amazon Comprehend.

Téléchargement de la sortie dans le compartiment S3

Après avoir téléchargé et extrait le fichier d'analyse des entités Amazon Comprehend, vous chargez le fichier `output` extrait dans votre compartiment Amazon S3.

Pour télécharger le fichier de sortie Amazon Comprehend extrait (console)

1. Ouvrez la console Amazon S3 à l'adresse https://console.aws.amazon.com/s3/.
2. Dans Buckets, cliquez sur le nom de votre bucket, puis choisissez Upload.
3. Dans Fichiers et dossiers, choisissez Ajouter des fichiers.
4. Dans la boîte de dialogue, accédez au `output` fichier extrait sur votre appareil, sélectionnez-le, puis choisissez Ouvrir.
5. Conservez les paramètres par défaut pour la destination, les autorisations et les propriétés.
6. Choisissez Charger.

Pour télécharger le fichier de sortie Amazon Comprehend extrait ()AWS CLI

1. Pour télécharger le `output` fichier extrait dans votre bucket, utilisez la commande copy :

Linux

```
aws s3 cp path/output s3://amzn-s3-demo-bucket/output
```

Où :

- *path/*est le chemin de fichier local vers votre fichier extraitoutput,
- amzn-s3-demo-bucket est le nom de votre compartiment S3.

macOS

```
aws s3 cp path/output s3://amzn-s3-demo-bucket/output
```

Où :

- *path/*est le chemin de fichier local vers votre fichier extraitoutput,
- amzn-s3-demo-bucket est le nom de votre compartiment S3.

Windows

```
aws s3 cp path/output s3://amzn-s3-demo-bucket/output
```

Où :

- *path/*est le chemin de fichier local vers votre fichier extraitoutput,
- amzn-s3-demo-bucket est le nom de votre compartiment S3.

2. Pour vous assurer que le `output` fichier a bien été chargé dans votre compartiment S3, vérifiez son contenu à l'aide de la commande list :

Linux

```
aws s3 ls s3://amzn-s3-demo-bucket/
```

Où :

- amzn-s3-demo-bucket est le nom de votre compartiment S3.

macOS

```
aws s3 ls s3://amzn-s3-demo-bucket/
```

Où :

- amzn-s3-demo-bucket est le nom de votre compartiment S3.

Windows

```
aws s3 ls s3://amzn-s3-demo-bucket/
```

Où :

- amzn-s3-demo-bucket est le nom de votre compartiment S3.

Conversion de la sortie au format de métadonnées Amazon Kendra

Pour convertir la sortie Amazon Comprehend en métadonnées Amazon Kendra, vous devez exécuter un script Python 3. Si vous utilisez la console, vous devez l'utiliser AWS CloudShell pour cette étape.

Pour exécuter le script Python 3 (console)

1. Téléchargez le fichier compressé .py.zip du convertisseur sur votre appareil.
2. Extrayez le fichier Python `3converter.py`.
3. Connectez-vous à la console AWS de gestion et assurez-vous que votre AWS région est définie sur la même région que votre compartiment S3 et votre tâche d'analyse Amazon Comprehend.
4. Cliquez sur l'AWS CloudShell icône ou saisissez du AWS CloudShelltexte dans la zone de recherche de la barre de navigation supérieure pour lancer un environnement.

> **Note**
>
> Lors du premier AWS CloudShell lancement dans une nouvelle fenêtre de navigateur, un panneau de bienvenue s'affiche et répertorie les principales fonctionnalités. Le shell est

prêt à interagir une fois que vous avez fermé ce panneau et que l'invite de commande s'affiche.

5. Une fois le terminal préparé, choisissez Actions dans le volet de navigation, puis choisissez Télécharger le fichier dans le menu.
6. Dans la boîte de dialogue qui s'ouvre, choisissez Sélectionner un fichier, puis choisissez le fichier `converter.py` Python 3 téléchargé sur votre appareil. Choisissez Charger.
7. Dans l' AWS CloudShell environnement, entrez la commande suivante :

```
python3 converter.py
```

8. Lorsque l'interface shell vous invite à saisir le nom de votre compartiment S3, entrez le nom de votre compartiment S3 et appuyez sur Entrée.
9. Lorsque l'interface shell vous invite à entrer le chemin de fichier complet vers votre fichier de sortie Comprehend, entrez et appuyez sur Entrée. **`output`**
10. Lorsque l'interface shell vous invite à saisir le chemin de fichier complet de votre dossier de métadonnées, entrez **`metadata/`** et appuyez sur Entrée.

> **Important**
>
> Pour que les métadonnées soient correctement formatées, les valeurs d'entrée des étapes 8 à 10 doivent être exactes.

Pour exécuter le script Python 3 (AWS CLI)

1. Pour télécharger le fichier Python 3`converter.py`, exécutez la commande suivante dans une fenêtre de terminal :

   Linux

   ```
   curl -o path/converter.py.zip https://docs.aws.amazon.com/kendra/latest/dg/
   samples/converter.py.zip
   ```

   Où :

- *path/*est le chemin du fichier vers l'emplacement dans lequel vous souhaitez enregistrer le fichier compressé.

macOS

```
curl -o path/converter.py.zip https://docs.aws.amazon.com/kendra/latest/dg/
samples/converter.py.zip
```

Où :

- *path/*est le chemin du fichier vers l'emplacement dans lequel vous souhaitez enregistrer le fichier compressé.

Windows

```
curl -o path/converter.py.zip https://docs.aws.amazon.com/kendra/latest/dg/
samples/converter.py.zip
```

Où :

- *path/*est le chemin du fichier vers l'emplacement dans lequel vous souhaitez enregistrer le fichier compressé.

2. Pour extraire le fichier Python 3, exécutez la commande suivante dans la fenêtre du terminal :

Linux

```
unzip path/converter.py.zip -d path/
```

Où :

- *path/*est le chemin du fichier enregistré. `converter.py.zip`

macOS

```
unzip path/converter.py.zip -d path/
```

Où :

- *path/*est le chemin du fichier enregistré. `converter.py.zip`

Windows

```
tar -xf path/converter.py.zip -C path/
```

Où :

- *path/*est le chemin du fichier enregistré. `converter.py.zip`

3. Assurez-vous que Boto3 est installé sur votre appareil en exécutant la commande suivante.

   Linux

   ```
   pip3 show boto3
   ```

   macOS

   ```
   pip3 show boto3
   ```

   Windows

   ```
   pip3 show boto3
   ```

   > **Note**
   >
   > Si Boto3 n'est pas installé, lancez-vous `pip3 install boto3` pour l'installer.

4. Pour exécuter le script Python 3 afin de convertir le `output` fichier, exécutez la commande suivante.

   Linux

   ```
   python path/converter.py
   ```

Où :

- *path/*est le chemin du fichier enregistré. `converter.py.zip`

macOS

```
python path/converter.py
```

Où :

- *path/*est le chemin du fichier enregistré. `converter.py.zip`

Windows

```
python path/converter.py
```

Où :

- *path/*est le chemin du fichier enregistré. `converter.py.zip`

5. Lorsque vous y AWS CLI êtes invité`Enter the name of your S3 bucket`, entrez le nom de votre compartiment S3 et appuyez sur Entrée.
6. Lorsque vous y AWS CLI êtes invité`Enter the full filepath to your Comprehend output file`, entrez **`output`** et appuyez sur Entrée.
7. Lorsque vous y AWS CLI êtes invité`Enter the full filepath to your metadata folder`, entrez **`metadata/`** et appuyez sur Entrée.

> **Important**
>
> Pour que les métadonnées soient correctement formatées, les valeurs d'entrée des étapes 5 à 7 doivent être exactes.

À la fin de cette étape, les métadonnées formatées sont déposées dans le `metadata` dossier de votre compartiment S3.

Nettoyage de votre compartiment Amazon S3

Étant donné que l'index Amazon Kendra synchronise tous les fichiers stockés dans un compartiment, nous vous recommandons de nettoyer votre compartiment Amazon S3 pour éviter les résultats de recherche redondants.

Pour nettoyer votre compartiment Amazon S3 (console)

1. Ouvrez la console Amazon S3 à l'adresse https://console.aws.amazon.com/s3/.
2. Dans Buckets, choisissez votre compartiment, puis sélectionnez le dossier de sortie de l'analyse des entités Amazon Comprehend, le fichier d'analyse des entités `.temp` Amazon Comprehend et le fichier Amazon Comprehend extrait. `output`
3. Dans l'onglet Vue d'ensemble, choisissez Supprimer.
4. Dans Supprimer des objets, sélectionnez Supprimer définitivement des objets ? et entrez **`permanently delete`** dans le champ de saisie de texte.
5. Choisissez Supprimer les objets.

Pour nettoyer votre compartiment Amazon S3 (AWS CLI)

1. Pour supprimer tous les fichiers et dossiers de votre compartiment S3 à l'exception `metadata` des dossiers `data` et, utilisez la commande remove dans le AWS CLI :

   Linux

   ```
   aws s3 rm s3://amzn-s3-demo-bucket/ --recursive --exclude "data/*" --exclude
    "metadata/*"
   ```

   Où :

   - amzn-s3-demo-bucket est le nom de votre compartiment S3.

   macOS

   ```
   aws s3 rm s3://amzn-s3-demo-bucket/ --recursive --exclude "data/*" --exclude
    "metadata/*"
   ```

   Où :

- amzn-s3-demo-bucket est le nom de votre compartiment S3.

Windows

```
aws s3 rm s3://amzn-s3-demo-bucket/ --recursive --exclude "data/*" --exclude
 "metadata/*"
```

Où :

- amzn-s3-demo-bucket est le nom de votre compartiment S3.

2. Pour vous assurer que les objets ont bien été supprimés de votre compartiment S3, vérifiez son contenu à l'aide de la commande list :

Linux

```
aws s3 ls s3://amzn-s3-demo-bucket/
```

Où :

- amzn-s3-demo-bucket est le nom de votre compartiment S3.

macOS

```
aws s3 ls s3://amzn-s3-demo-bucket/
```

Où :

- amzn-s3-demo-bucket est le nom de votre compartiment S3.

Windows

```
aws s3 ls s3://amzn-s3-demo-bucket/
```

Où :

- amzn-s3-demo-bucket est le nom de votre compartiment S3.

À la fin de cette étape, vous avez converti les résultats de l'analyse des entités Amazon Comprehend en métadonnées Amazon Kendra. Vous êtes maintenant prêt à créer un index Amazon Kendra.

Étape 4 : Création d'un index Amazon Kendra et ingestion des métadonnées

Pour implémenter votre solution de recherche intelligente, vous devez créer un index Amazon Kendra et y ingérer vos données et métadonnées S3.

Avant d'ajouter des métadonnées à votre index Amazon Kendra, vous devez créer des champs d'index personnalisés correspondant à des attributs de document personnalisés, qui à leur tour correspondent aux types d'entités Amazon Comprehend. Amazon Kendra utilise les champs d'index et les attributs de document personnalisés que vous créez pour rechercher et filtrer vos documents.

Pour plus d'informations, voir Index et Création d'attributs de document personnalisés.

Rubriques

- Création d'un index Amazon Kendra
- Mise à jour du rôle IAM pour l'accès à Amazon S3
- Création de champs d'index de recherche personnalisés pour Amazon Kendra
- Ajout du compartiment Amazon S3 en tant que source de données pour l'index
- Synchronisation de l'index Amazon Kendra

Création d'un index Amazon Kendra

Pour interroger vos documents sources, vous devez créer un index Amazon Kendra.

Si vous utilisez le AWS CLI dans cette étape, vous créez et attachez un rôle et une politique AWS IAM qui permettent à Amazon Kendra d'accéder à CloudWatch vos journaux avant de créer un index. Pour plus d'informations, veuillez consulter les Prérequis.

Pour créer un index Amazon Kendra (console)

1. Ouvrez la console Amazon Kendra à l'adresse. https://console.aws.amazon.com/kendra/

> **Important**
>
> Assurez-vous que vous vous trouvez dans la même région que celle dans laquelle vous avez créé votre tâche d'analyse des entités Amazon Comprehend et votre compartiment Amazon S3. Si vous vous trouvez dans une autre région, choisissez la AWS région dans laquelle vous avez créé votre compartiment Amazon S3 dans le sélecteur de région situé dans la barre de navigation supérieure.

2. Choisissez Créer un index.
3. Pour les détails de l'index sur la page Spécifier les détails de l'index, procédez comme suit :
   a. Pour Nom de l'index, saisissez **`kendra-index`**.
   b. Laissez le champ Description vide.
   c. Pour Rôle IAM, choisissez Créer un rôle. Ce rôle permet d'accéder à votre compartiment Amazon S3.
   d. Pour le Nom du rôle, saisissez **`kendra-role`**. Le rôle IAM aura le préfixe`AmazonKendra-`.
   e. Conservez les paramètres par défaut pour le chiffrement et les balises, puis choisissez Next.
4. Pour les paramètres de contrôle d'accès sur la page Configurer le contrôle d'accès utilisateur, choisissez Non, puis Suivant.
5. Pour les éditions Provisioning sur la page des détails de Provisioning, choisissez Developer edition puis Create.

Pour créer un index Amazon Kendra ()AWS CLI

1. Pour créer et associer un rôle IAM à Amazon Kendra qui le reconnaisse comme une entité de confiance, procédez comme suit :
   a. Enregistrez la politique de confiance suivante sous forme de fichier JSON appelé `kendra-trust-policy.json` dans un éditeur de texte sur votre appareil local.

      JSON

```
{
    "Version": "2012-10-17",
    "Statement": {
```

```
            "Effect": "Allow",
            "Principal": {
                "Service": "kendra.amazonaws.com"
            },
            "Action": "sts:AssumeRole"
        }
}
```

b. Pour créer un rôle IAM appelé `kendra-role` et y joindre votre `kendra-trust-policy.json` fichier enregistré, utilisez la commande create-role :

Linux

```
aws iam create-role \
          --role-name kendra-role \
          --assume-role-policy-document file://path/kendra-trust-policy.json
```

Où :

- *path/*est le chemin du fichier vers votre `kendra-trust-policy.json` appareil local.

macOS

```
aws iam create-role \
          --role-name kendra-role \
          --assume-role-policy-document file://path/kendra-trust-policy.json
```

Où :

- *path/*est le chemin du fichier vers votre `kendra-trust-policy.json` appareil local.

Windows

```
aws iam create-role ^
          --role-name kendra-role ^
          --assume-role-policy-document file://path/kendra-trust-policy.json
```

Où :

- *path/*est le chemin du fichier vers votre `kendra-trust-policy.json` appareil local.

c. Copiez l'Amazon Resource Name (ARN) dans votre éditeur de texte et enregistrez-le localement sous le nom de`kendra-role-arn`.

> **Note**
>
> Le format de l'ARN est similaire à*arn:aws:iam::123456789012:role/kendra-role*. Vous avez besoin de l'ARN sous lequel vous avez enregistré `kendra-role-arn` pour exécuter les tâches Amazon Kendra.

2. Avant de créer un index, vous devez autoriser `kendra-role` l'écriture dans CloudWatch Logs. Pour y arriver, exécutez les étapes suivantes.

a. Enregistrez la politique de confiance suivante sous forme de fichier JSON appelé `kendra-cloudwatch-policy.json` dans un éditeur de texte sur votre appareil local.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Action":"cloudwatch:PutMetricData",
         "Resource":"*",
         "Condition":{
            "StringEquals":{
               "cloudwatch:namespace":"Kendra"
            }
         }
      },
      {
         "Effect":"Allow",
         "Action":"logs:DescribeLogGroups",
         "Resource":"*"
      },
```

```
      {
         "Effect":"Allow",
         "Action":"logs:CreateLogGroup",
         "Resource":"arn:aws:logs:aws-region:aws-account-id:log-group:/
aws/kendra/*"
      },
      {
         "Effect":"Allow",
         "Action":[
            "logs:DescribeLogStreams",
            "logs:CreateLogStream",
            "logs:PutLogEvents"
         ],
         "Resource":"arn:aws:logs:aws-region:aws-account-id:log-group:/
aws/kendra/*:log-stream:*"
      }
   ]
}
```

Remplacez *aws-region* par votre AWS région et *aws-account-id* par votre identifiant de AWS compte à 12 chiffres.

b. Pour créer une politique IAM permettant d'accéder aux CloudWatch journaux, utilisez la commande create-policy :

Linux

```
aws iam create-policy \
         --policy-name kendra-cloudwatch-policy \
         --policy-document file://path/kendra-cloudwatch-policy.json
```

Où :

- *path/*est le chemin du fichier vers votre `kendra-cloudwatch-policy.json` appareil local.

macOS

```
aws iam create-policy \
         --policy-name kendra-cloudwatch-policy \
```

```
        --policy-document file://path/kendra-cloudwatch-policy.json
```

Où :

- *path/*est le chemin du fichier vers votre `kendra-cloudwatch-policy.json` appareil local.

Windows

```
aws iam create-policy ^
        --policy-name kendra-cloudwatch-policy ^
        --policy-document file://path/kendra-cloudwatch-policy.json
```

Où :

- *path/*est le chemin du fichier vers votre `kendra-cloudwatch-policy.json` appareil local.

c. Copiez l'Amazon Resource Name (ARN) dans votre éditeur de texte et enregistrez-le localement sous le nom de`kendra-cloudwatch-arn`.

> **Note**
>
> Le format de l'ARN est similaire à*arn:aws:iam::123456789012:role/kendra-cloudwatch-policy*. Vous avez besoin de l'ARN sous lequel vous avez enregistré `kendra-cloudwatch-arn` pour l'associer `kendra-cloudwatch-policy` à votre rôle IAM.

d. Pour l'associer `kendra-cloudwatch-policy` à votre rôle IAM, utilisez la attach-role-policycommande :

Linux

```
aws iam attach-role-policy \
        --policy-arn policy-arn \
        --role-name kendra-role
```

Où :

- *policy-arn*est votre sauvegardekendra-cloudwatch-arn.

macOS

```
aws iam attach-role-policy \
          --policy-arn policy-arn \
          --role-name kendra-role
```

Où :

- *policy-arn*est votre sauvegardekendra-cloudwatch-arn.

Windows

```
aws iam attach-role-policy ^
          --policy-arn policy-arn ^
          --role-name kendra-role
```

Où :

- *policy-arn*est votre sauvegardekendra-cloudwatch-arn.

3. Pour créer un index, utilisez la commande create-index :

Linux

```
aws kendra create-index \
        --name kendra-index \
        --edition DEVELOPER_EDITION \
        --role-arn role-arn \
        --region aws-region
```

Où :

- *role-arn*est votre sauvegardekendra-role-arn,
- *aws-region*est votre AWS région.

macOS

```
aws kendra create-index \
        --name kendra-index \
        --edition DEVELOPER_EDITION \
        --role-arn role-arn \
        --region aws-region
```

Où :

- *role-arn*est votre sauvegardekendra-role-arn,
- *aws-region*est votre AWS région.

Windows

```
aws kendra create-index ^
        --name kendra-index ^
        --edition DEVELOPER_EDITION ^
        --role-arn role-arn ^
        --region aws-region
```

Où :

- *role-arn*est votre sauvegardekendra-role-arn,
- *aws-region*est votre AWS région.

4. Copiez l'index Id et enregistrez-le dans un éditeur de texte souskendra-index-id. IdCela vous permet de suivre l'état de la création de votre index.
5. Pour suivre la progression de votre tâche de création d'index, utilisez la commande describe-index :

Linux

```
aws kendra describe-index \
        --id kendra-index-id \
        --region aws-region
```

Où :

- *kendra-index-id*est votre sauvegardekendra-index-id,
- *aws-region*est votre AWS région.

macOS

```
aws kendra describe-index \
        --id kendra-index-id \
        --region aws-region
```

Où :

- *kendra-index-id*est votre sauvegardekendra-index-id,
- *aws-region*est votre AWS région.

Windows

```
aws kendra describe-index ^
        --id kendra-index-id ^
        --region aws-region
```

Où :

- *kendra-index-id*est votre sauvegardekendra-index-id,
- *aws-region*est votre AWS région.

Le processus de création d'index prend en moyenne 15 minutes, mais peut prendre plus de temps. Lorsque le statut de l'index est actif, votre index est prêt à être utilisé. Pendant la création de votre index, vous pouvez passer à l'étape suivante.

Si vous utilisez le AWS CLI dans cette étape, vous créez et associez une politique IAM à votre rôle Amazon Kendra IAM qui donne à votre index les autorisations d'accéder à votre compartiment S3.

Mise à jour du rôle IAM pour l'accès à Amazon S3

Pendant la création de l'index, vous mettez à jour votre rôle Amazon Kendra IAM pour permettre à l'index que vous avez créé de lire les données de votre compartiment Amazon S3. Pour plus d'informations, consultez la section Rôles d'accès IAM pour Amazon Kendra.

Pour mettre à jour votre rôle IAM (console)

1. Ouvrez la console IAM à l'adresse https://console.aws.amazon.com/iam/.
2. Dans le volet de navigation de gauche, choisissez Rôles et entrez **`kendra-role`** dans la zone de recherche au-dessus du nom du rôle.
3. Parmi les options proposées, cliquez sur`kendra-role`.
4. Dans Résumé, choisissez Joindre des politiques.
5. Dans Autorisations d'attachement, dans la zone de recherche, entrez **`S3`** et cochez la case à côté de la ReadOnlyAccess politique AmazonS3 parmi les options suggérées.
6. Choisissez Attach policy (Attacher la politique). Sur la page Résumé, vous pouvez désormais voir deux politiques associées au rôle IAM.
7. Revenez à la console Amazon Kendra à l'adresse https://console.aws.amazon.com/kendra/et attendez que le statut de votre index passe de Création à Actif avant de passer à l'étape suivante.

Pour mettre à jour votre rôle IAM ()AWS CLI

1. Enregistrez le texte suivant dans un fichier JSON appelé `kendra-S3-access-policy.json` dans un éditeur de texte sur votre appareil local.

   JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Action":[
            "s3:GetObject"
         ],
         "Resource":[
            "arn:aws:s3:::amzn-s3-demo-bucket/*"
         ],
         "Effect":"Allow"
      },
      {
         "Action":[
            "s3:ListBucket"
         ],
```

```
            "Resource":[
               "arn:aws:s3:::amzn-s3-demo-bucket"
            ],
            "Effect":"Allow"
         },
         {
            "Effect":"Allow",
            "Action":[
               "kendra:BatchPutDocument",
               "kendra:BatchDeleteDocument",
               "kendra:ListDataSourceSyncJobs"
            ],
            "Resource":[
               "arn:aws:kendra:aws-region:aws-account-id:index/kendra-index-id"
            ]
         }
      ]
}
```

Remplacez amzn-s3-demo-bucket par le nom de votre compartiment S3, par votre AWS région, *aws-region* par votre identifiant de compte à 12 chiffres et *aws-account-id* par votre sauvegarde AWS . *kendra-index-id* kendra-index-id

2. Pour créer une politique IAM permettant d'accéder à votre compartiment S3, utilisez la commande create-policy :

   Linux

   ```
   aws iam create-policy \
            --policy-name kendra-S3-access-policy \
            --policy-document file://path/kendra-S3-access-policy.json
   ```

   Où :

   - *path/*est le chemin du fichier vers votre kendra-S3-access-policy.json appareil local.

   macOS

   ```
   aws iam create-policy \
   ```

```
        --policy-name kendra-S3-access-policy \
        --policy-document file://path/kendra-S3-access-policy.json
```

Où :

- *path/*est le chemin du fichier vers votre `kendra-S3-access-policy.json` appareil local.

Windows

```
aws iam create-policy ^
        --policy-name kendra-S3-access-policy ^
        --policy-document file://path/kendra-S3-access-policy.json
```

Où :

- *path/*est le chemin du fichier vers votre `kendra-S3-access-policy.json` appareil local.

3. Copiez l'Amazon Resource Name (ARN) dans votre éditeur de texte et enregistrez-le localement sous le nom de`kendra-S3-access-arn`.

> **Note**
>
> Le format de l'ARN est similaire à*arn:aws:iam::123456789012:role/kendra-S3-access-policy*. Vous avez besoin de l'ARN sous lequel vous avez enregistré `kendra-S3-access-arn` pour l'associer `kendra-S3-access-policy` à votre rôle IAM.

4. Pour l'associer `kendra-S3-access-policy` à votre rôle Amazon Kendra IAM, utilisez la commande suivante : attach-role-policy

Linux

```
aws iam attach-role-policy \
        --policy-arn policy-arn \
        --role-name kendra-role
```

Où :

- *policy-arn*est votre sauvegardekendra-S3-access-arn.

macOS

```
aws iam attach-role-policy \
          --policy-arn policy-arn \
          --role-name kendra-role
```

Où :

- *policy-arn*est votre sauvegardekendra-S3-access-arn.

Windows

```
aws iam attach-role-policy ^
          --policy-arn policy-arn ^
          --role-name kendra-role
```

Où :

- *policy-arn*est votre sauvegardekendra-S3-access-arn.

Création de champs d'index de recherche personnalisés pour Amazon Kendra

Pour préparer Amazon Kendra à reconnaître vos métadonnées comme des attributs de document personnalisés, vous devez créer des champs personnalisés correspondant aux types d'entités Amazon Comprehend. Vous saisissez les neuf types d'entités Amazon Comprehend suivants sous forme de champs personnalisés :

- ARTICLE_COMMERCIAL
- DATE
- ÉVÉNEMENT
- LOCATION
- ORGANISATION

- OTHER
- LA PERSONNE
- QUANTITÉ
- TITLE

> **Important**
>
> Les types d'entités mal orthographiés ne seront pas reconnus par l'index.

Pour créer des champs personnalisés pour votre index Amazon Kendra (console)

1. Ouvrez la console Amazon Kendra à l'adresse. https://console.aws.amazon.com/kendra/
2. Dans la liste des index, cliquez sur`kendra-index`.
3. Dans le panneau de navigation de gauche, sous Gestion des données, choisissez Définition des facettes.
4. Dans le menu des champs d'index, choisissez Ajouter un champ.
5. Dans la boîte de dialogue Ajouter un champ d'index, procédez comme suit :
   a. Dans Nom du champ, entrez**`COMMERCIAL_ITEM`**.
   b. Dans Type de données, sélectionnez Liste de chaînes.
   c. Dans Types d'utilisation, sélectionnez Facetable, Consultable et Displayable, puis choisissez Ajouter.
   d. Répétez les étapes a à c pour chaque type d'entité Amazon Comprehend : COMMERCIAL_ITEM, DATE, ÉVÉNEMENT, LIEU, ORGANISATION, AUTRE, PERSONNE, QUANTITÉ, TITRE.

La console affiche les messages d'ajout de champs réussis. Vous pouvez choisir de les fermer avant de passer à l'étape suivante.

Pour créer des champs personnalisés pour votre index Amazon Kendra ()AWS CLI

1. Enregistrez le texte suivant sous forme de fichier JSON appelé `custom-attributes.json` dans un éditeur de texte sur votre appareil local.

```
[
```

```
    {
        "Name": "COMMERCIAL_ITEM",
        "Type": "STRING_LIST_VALUE",
        "Search": {
            "Facetable": true,
            "Searchable": true,
            "Displayable": true
        }
    },
    {
        "Name": "DATE",
        "Type": "STRING_LIST_VALUE",
        "Search": {
            "Facetable": true,
            "Searchable": true,
            "Displayable": true
        }
    },
    {
        "Name": "EVENT",
        "Type": "STRING_LIST_VALUE",
        "Search": {
            "Facetable": true,
            "Searchable": true,
            "Displayable": true
        }
    },
    {
        "Name": "LOCATION",
        "Type": "STRING_LIST_VALUE",
        "Search": {
            "Facetable": true,
            "Searchable": true,
            "Displayable": true
        }
    },
    {
        "Name": "ORGANIZATION",
        "Type": "STRING_LIST_VALUE",
        "Search": {
            "Facetable": true,
            "Searchable": true,
            "Displayable": true
        }
```

```
    },
    {
        "Name": "OTHER",
        "Type": "STRING_LIST_VALUE",
        "Search": {
            "Facetable": true,
            "Searchable": true,
            "Displayable": true
        }
    },
    {
        "Name": "PERSON",
        "Type": "STRING_LIST_VALUE",
        "Search": {
            "Facetable": true,
            "Searchable": true,
            "Displayable": true
        }
    },
    {
        "Name": "QUANTITY",
        "Type": "STRING_LIST_VALUE",
        "Search": {
            "Facetable": true,
            "Searchable": true,
            "Displayable": true
        }
    },
    {
        "Name": "TITLE",
        "Type": "STRING_LIST_VALUE",
        "Search": {
            "Facetable": true,
            "Searchable": true,
            "Displayable": true
        }
    }
]
```

2. Pour créer des champs personnalisés dans votre index, utilisez la commande update-index :

   Linux

```
aws kendra update-index \
```

```
        --id kendra-index-id \
        --document-metadata-configuration-updates file://path/custom-
attributes.json \
        --region aws-region
```

Où :

- *kendra-index-id*est votre sauvegardekendra-index-id,
- *path/*est le chemin du fichier vers votre custom-attributes.json appareil local,
- *aws-region*est votre AWS région.

macOS

```
aws kendra update-index \
        --id kendra-index-id \
        --document-metadata-configuration-updates file://path/custom-
attributes.json \
        --region aws-region
```

Où :

- *kendra-index-id*est votre sauvegardekendra-index-id,
- *path/*est le chemin du fichier vers votre custom-attributes.json appareil local,
- *aws-region*est votre AWS région.

Windows

```
aws kendra update-index ^
        --id kendra-index-id ^
        --document-metadata-configuration-updates file://path/custom-
attributes.json ^
        --region aws-region
```

Où :

- *kendra-index-id*est votre sauvegardekendra-index-id,
- *path/*est le chemin du fichier vers votre custom-attributes.json appareil local,

- *aws-region*est votre AWS région.

3. Pour vérifier que les attributs personnalisés ont été ajoutés à votre index, utilisez la commande describe-index :

   Linux

   ```
   aws kendra describe-index \
           --id kendra-index-id \
           --region aws-region
   ```

   Où :

   - *kendra-index-id*est votre sauvegardekendra-index-id,
   - *aws-region*est votre AWS région.

   macOS

   ```
   aws kendra describe-index \
           --id kendra-index-id \
           --region aws-region
   ```

   Où :

   - *kendra-index-id*est votre sauvegardekendra-index-id,
   - *aws-region*est votre AWS région.

   Windows

   ```
   aws kendra describe-index ^
           --id kendra-index-id ^
           --region aws-region
   ```

   Où :

   - *kendra-index-id*est votre sauvegardekendra-index-id,
   - *aws-region*est votre AWS région.

Ajout du compartiment Amazon S3 en tant que source de données pour l'index

Avant de synchroniser votre index, vous devez y connecter votre source de données S3.

Pour connecter un compartiment S3 à votre index Amazon Kendra (console)

1. Ouvrez la console Amazon Kendra à l'adresse. https://console.aws.amazon.com/kendra/
2. Dans la liste des index, cliquez sur`kendra-index`.
3. Dans le menu de navigation de gauche, sous Gestion des données, sélectionnez Sources de données.
4. Dans la section Sélectionner le type de connecteur de source de données, accédez à Amazon S3 et choisissez Ajouter un connecteur.
5. Dans la page Spécifier les détails de la source de données, procédez comme suit :
   a. Sous Nom et description, dans Nom de la source de données, entrez**`S3-data-source`**.
   b. Laissez la section Description vide.
   c. Conservez les paramètres par défaut pour les balises.
   d. Choisissez Suivant.
6. Sur la page Configurer les paramètres de synchronisation, dans la section Étendue de la synchronisation, procédez comme suit :
   a. Dans Entrez l'emplacement de la source de données, choisissez Parcourir S3.
   b. Dans Choisir les ressources, sélectionnez votre compartiment S3, puis choisissez Choose.
   c. Dans Emplacement du dossier de préfixes des fichiers de métadonnées, choisissez Parcourir S3.
   d. Dans Choisir les ressources, cliquez sur le nom de votre bucket dans la liste des buckets.
   e. Pour Objets, sélectionnez la case d'option correspondant à `metadata` et choisissez Choisir. Le champ de localisation doit maintenant indiquer`metadata/`.
   f. Conservez les paramètres par défaut pour l'emplacement du fichier de configuration de la liste de contrôle d'accès, la sélection de la clé de déchiffrement et la configuration supplémentaire.
7. Pour le rôle IAM, sur la page Configurer les paramètres de synchronisation, sélectionnez`kendra-role`.

8. Sur la page Configurer les paramètres de synchronisation, sous Calendrier d'exécution de synchronisation, pour Fréquence, choisissez Exécuter à la demande, puis Suivant.
9. Sur la page Révision et création, passez en revue vos choix concernant les détails de la source de données et choisissez Ajouter une source de données.

Pour connecter un compartiment S3 à votre index Amazon Kendra ()AWS CLI

1. Enregistrez le texte suivant sous forme de fichier JSON appelé `S3-data-connector.json` dans un éditeur de texte sur votre appareil local.

```
{
   "S3Configuration":{
      "BucketName":"amzn-s3-demo-bucket",
      "DocumentsMetadataConfiguration":{
         "S3Prefix":"metadata"
      }
   }
}
```

   Remplacez amzn-s3-demo-bucket par le nom de votre compartiment S3.

2. Pour connecter votre compartiment S3 à votre index, utilisez la create-data-sourcecommande suivante :

   Linux

```
aws kendra create-data-source \
        --index-id kendra-index-id \
        --name S3-data-source \
        --type S3 \
        --configuration file://path/S3-data-connector.json \
        --role-arn role-arn \
        --region aws-region
```

   Où :

   - *kendra-index-id*est votre sauvegardekendra-index-id,
   - *path/*est le chemin du fichier vers votre S3-data-connector.json appareil local,
   - *role-arn*est votre sauvegardekendra-role-arn,

- *aws-region*est votre AWS région.

macOS

```
aws kendra create-data-source \
        --index-id kendra-index-id \
        --name S3-data-source \
        --type S3 \
        --configuration file://path/S3-data-connector.json \
        --role-arn role-arn \
        --region aws-region
```

Où :

- *kendra-index-id*est votre sauvegardekendra-index-id,
- *path/*est le chemin du fichier vers votre S3-data-connector.json appareil local,
- *role-arn*est votre sauvegardekendra-role-arn,
- *aws-region*est votre AWS région.

Windows

```
aws kendra create-data-source ^
        --index-id kendra-index-id ^
        --name S3-data-source ^
        --type S3 ^
        --configuration file://path/S3-data-connector.json ^
        --role-arn role-arn ^
        --region aws-region
```

Où :

- *kendra-index-id*est votre sauvegardekendra-index-id,
- *path/*est le chemin du fichier vers votre S3-data-connector.json appareil local,
- *role-arn*est votre sauvegardekendra-role-arn,
- *aws-region*est votre AWS région.

3. Copiez le connecteur Id et enregistrez-le dans un éditeur de texte sousS3-connector-id. IdCela vous permet de suivre l'état du processus de connexion des données.

4. Pour vous assurer que votre source de données S3 a été correctement connectée, utilisez la describe-data-sourcecommande suivante :

   Linux

   ```
   aws kendra describe-data-source \
           --id S3-connector-id \
           --index-id kendra-index-id \
           --region aws-region
   ```

   Où :

   - *S3-connector-id*est votre sauvegarde`S3-connector-id`,
   - *kendra-index-id*est votre sauvegarde`kendra-index-id`,
   - *aws-region*est votre AWS région.

   macOS

   ```
   aws kendra describe-data-source \
           --id S3-connector-id \
           --index-id kendra-index-id \
           --region aws-region
   ```

   Où :

   - *S3-connector-id*est votre sauvegarde`S3-connector-id`,
   - *kendra-index-id*est votre sauvegarde`kendra-index-id`,
   - *aws-region*est votre AWS région.

   Windows

   ```
   aws kendra describe-data-source ^
           --id S3-connector-id ^
           --index-id kendra-index-id ^
           --region aws-region
   ```

   Où :

- *S3-connector-id*est votre sauvegardeS3-connector-id,
- *kendra-index-id*est votre sauvegardekendra-index-id,
- *aws-region*est votre AWS région.

À la fin de cette étape, votre source de données Amazon S3 est connectée à l'index.

Synchronisation de l'index Amazon Kendra

Avec l'ajout de la source de données Amazon S3, vous pouvez désormais synchroniser votre index Amazon Kendra avec celle-ci.

Pour synchroniser votre index Amazon Kendra (console)

1. Ouvrez la console Amazon Kendra à l'adresse. https://console.aws.amazon.com/kendra/
2. Dans la liste des index, cliquez surkendra-index.
3. Dans le menu de navigation de gauche, sélectionnez Sources de données.
4. Dans Sources de données, sélectionnezS3-data-source.
5. Dans la barre de navigation supérieure, choisissez Synchroniser maintenant.

Pour synchroniser votre index Amazon Kendra ()AWS CLI

1. Pour synchroniser votre index, utilisez la commande start-data-source-sync-job :

   Linux

   ```
   aws kendra start-data-source-sync-job \
           --id S3-connector-id \
           --index-id kendra-index-id \
           --region aws-region
   ```

   Où :

   - *S3-connector-id*est votre sauvegardeS3-connector-id,
   - *kendra-index-id*est votre sauvegardekendra-index-id,
   - *aws-region*est votre AWS région.

macOS

```
aws kendra start-data-source-sync-job \
        --id S3-connector-id \
        --index-id kendra-index-id \
        --region aws-region
```

Où :

- *S3-connector-id*est votre sauvegarde`S3-connector-id`,
- *kendra-index-id*est votre sauvegarde`kendra-index-id`,
- *aws-region*est votre AWS région.

Windows

```
aws kendra start-data-source-sync-job ^
        --id S3-connector-id ^
        --index-id kendra-index-id ^
        --region aws-region
```

Où :

- *S3-connector-id*est votre sauvegarde`S3-connector-id`,
- *kendra-index-id*est votre sauvegarde`kendra-index-id`,
- *aws-region*est votre AWS région.

2. Pour vérifier l'état de la synchronisation des index, utilisez la commande list-data-source-sync-jobs :

Linux

```
aws kendra list-data-source-sync-jobs \
        --id S3-connector-id \
        --index-id kendra-index-id \
        --region aws-region
```

Où :

- *S3-connector-id*est votre sauvegardeS3-connector-id,
- *kendra-index-id*est votre sauvegardekendra-index-id,
- *aws-region*est votre AWS région.

macOS

```
aws kendra list-data-source-sync-jobs \
        --id S3-connector-id \
        --index-id kendra-index-id \
        --region aws-region
```

Où :

- *S3-connector-id*est votre sauvegardeS3-connector-id,
- *kendra-index-id*est votre sauvegardekendra-index-id,
- *aws-region*est votre AWS région.

Windows

```
aws kendra list-data-source-sync-jobs ^
        --id S3-connector-id ^
        --index-id kendra-index-id ^
        --region aws-region
```

Où :

- *S3-connector-id*est votre sauvegardeS3-connector-id,
- *kendra-index-id*est votre sauvegardekendra-index-id,
- *aws-region*est votre AWS région.

À la fin de cette étape, vous avez créé un index Amazon Kendra consultable et filtrable pour votre ensemble de données.

Étape 5 : Interrogation de l'index Amazon Kendra

Votre index Amazon Kendra est désormais prêt pour les requêtes en langage naturel. Lorsque vous effectuez une recherche dans votre index, Amazon Kendra utilise toutes les données et métadonnées que vous avez fournies pour renvoyer les réponses les plus précises à votre requête de recherche.

Amazon Kendra peut répondre à trois types de requêtes :

- Requêtes factuelles (questions « qui », « quoi », « quand » ou « où »)
- Requêtes descriptives (questions « comment »)
- Recherches par mots clés (questions dont l'intention et la portée ne sont pas claires)

Rubriques

- Interrogation de votre index Amazon Kendra
- Filtrer les résultats de recherche

Interrogation de votre index Amazon Kendra

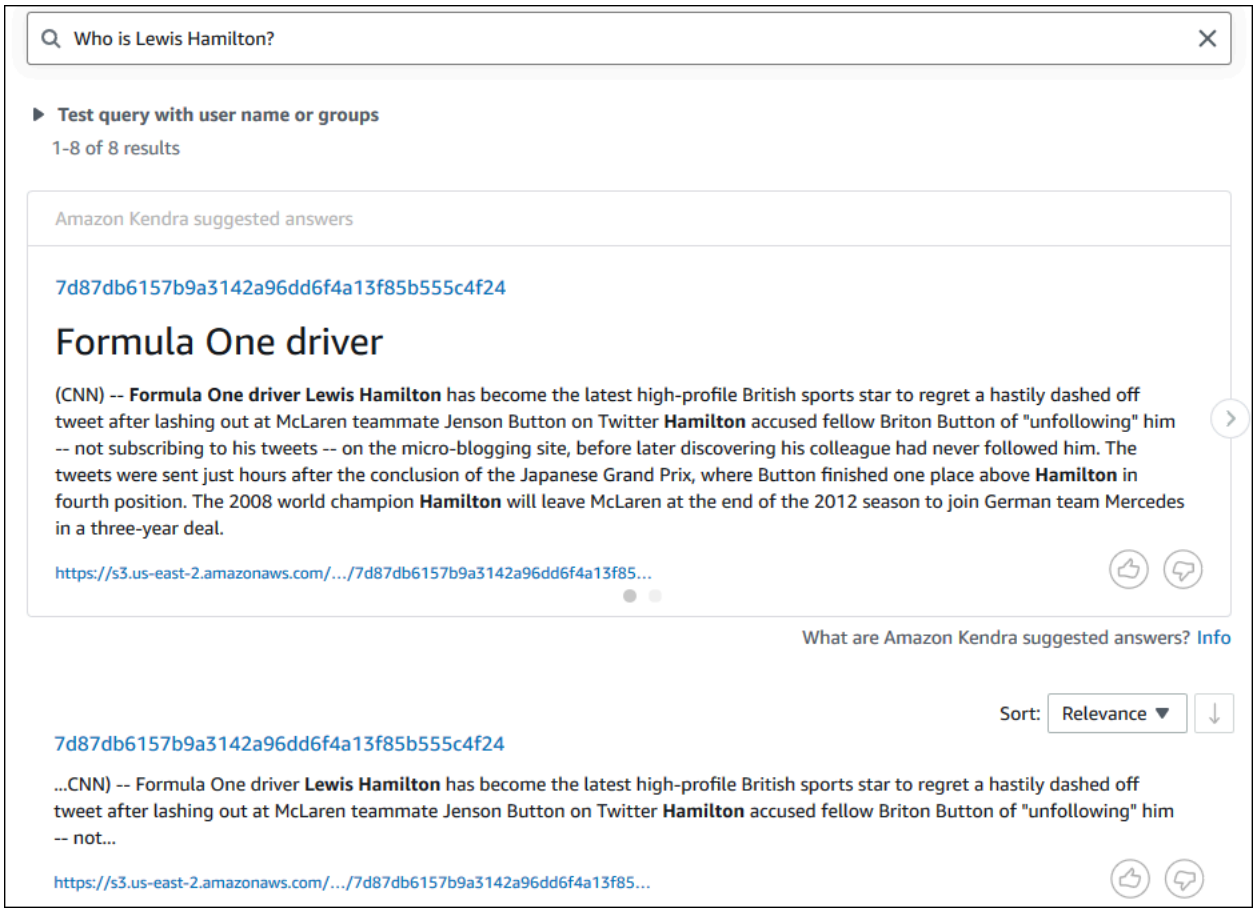
Vous pouvez interroger votre index Amazon Kendra à l'aide de questions correspondant aux trois types de requêtes pris en charge par Amazon Kendra. Pour plus d'informations, consultez la section Requêtes.

Les exemples de questions de cette section ont été choisis en fonction de l'ensemble de données d'exemple.

Pour interroger votre index Amazon Kendra (console)

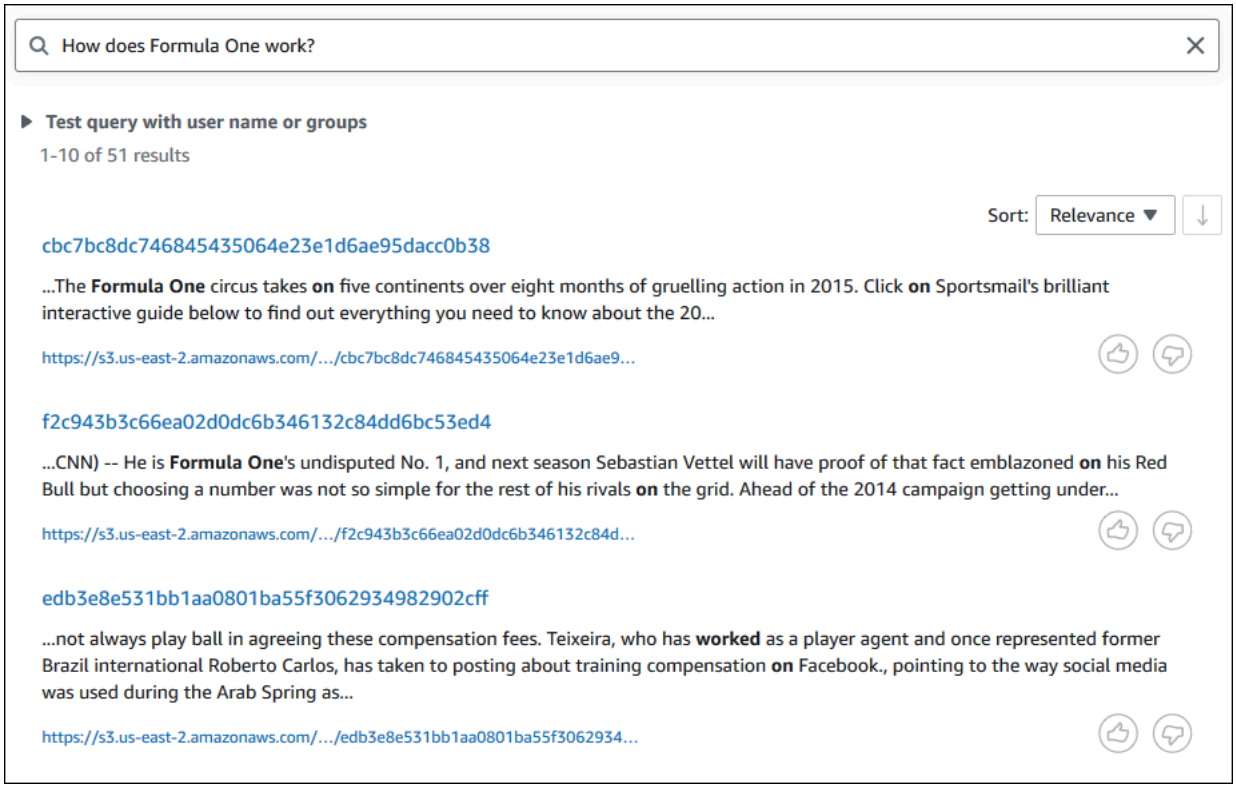
1. Ouvrez la console Amazon Kendra à l'adresse. https://console.aws.amazon.com/kendra/
2. Dans la liste des index, cliquez sur`kendra-index`.
3. Dans le menu de navigation de gauche, choisissez l'option de recherche dans votre index.
4. Pour exécuter un exemple de requête factoïdale, entrez **`Who is Lewis Hamilton?`** dans le champ de recherche et appuyez sur Entrée.

   Le premier résultat renvoyé est la réponse suggérée par Amazon Kendra, ainsi que le fichier de données contenant la réponse. Le reste des résultats constitue l'ensemble des documents recommandés.



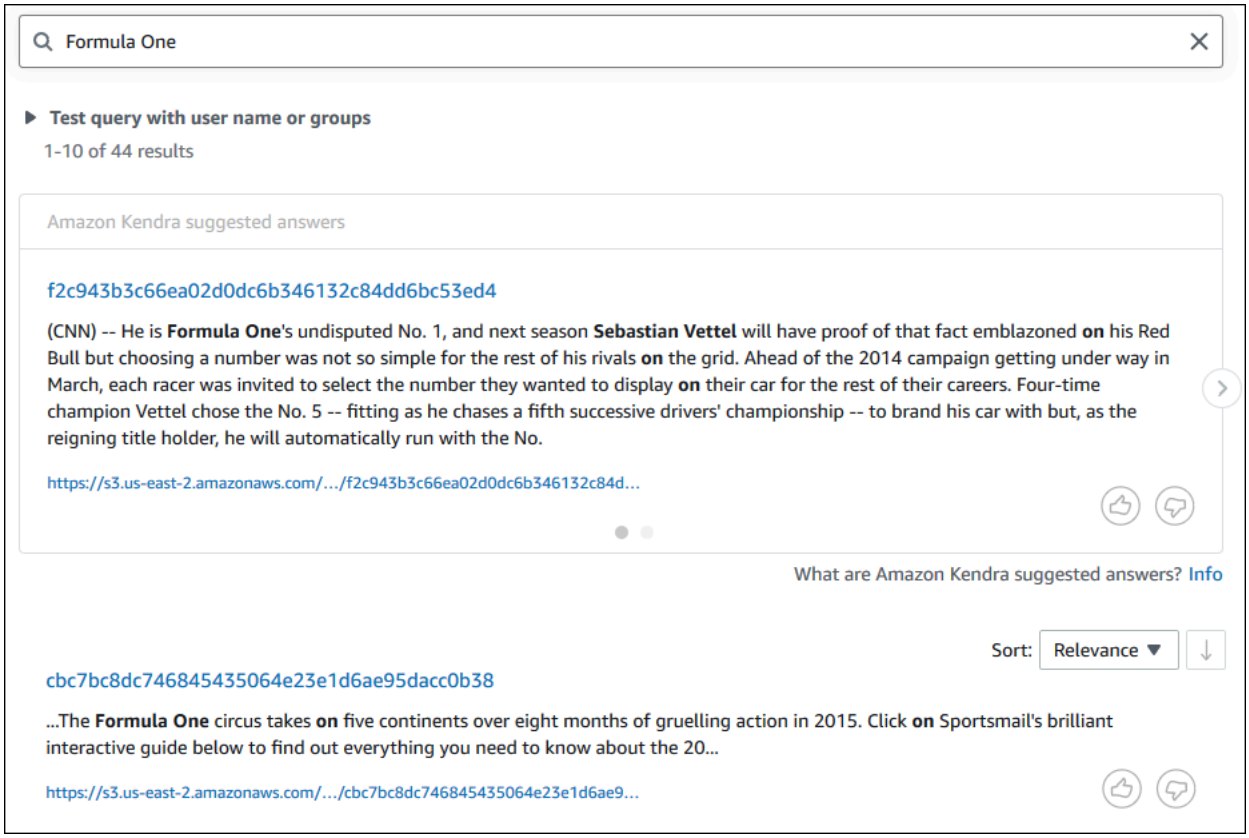
5. Pour exécuter une requête descriptive, entrez **`How does Formula One work?`** dans le champ de recherche et appuyez sur Entrée.

   Vous verrez un autre résultat renvoyé par la console Amazon Kendra, cette fois avec la phrase correspondante surlignée.



6. Pour lancer une recherche par mot clé, entrez **`Formula One`** dans le champ de recherche et appuyez sur Entrée.

   Vous verrez un autre résultat renvoyé par la console Amazon Kendra, suivi des résultats pour toutes les autres mentions de la phrase dans l'ensemble de données.



Pour interroger votre index Amazon Kendra ()AWS CLI

1. Pour exécuter un exemple de requête factoid, utilisez la commande query :

   Linux

   ```
   aws kendra query \
           --index-id kendra-index-id \
           --query-text "Who is Lewis Hamilton?" \
           --region aws-region
   ```

   Où :

   - *kendra-index-id*est votre sauvegardekendra-index-id,
   - *aws-region*est votre AWS région.

macOS

```
aws kendra query \
        --index-id kendra-index-id \
        --query-text "Who is Lewis Hamilton?" \
        --region aws-region
```

Où :

- *kendra-index-id*est votre sauvegardekendra-index-id,
- *aws-region*est votre AWS région.

Windows

```
aws kendra query ^
        --index-id kendra-index-id ^
        --query-text "Who is Lewis Hamilton?" ^
        --region aws-region
```

Où :

- *kendra-index-id*est votre sauvegardekendra-index-id,
- *aws-region*est votre AWS région.

AWS CLI Affiche les résultats de votre requête.

2. Pour exécuter un exemple de requête descriptive, utilisez la commande query :

Linux

```
aws kendra query \
        --index-id kendra-index-id \
        --query-text "How does Formula One work?" \
        --region aws-region
```

Où :

- *kendra-index-id*est votre sauvegardekendra-index-id,

- *aws-region*est votre AWS région.

macOS

```
aws kendra query \
        --index-id kendra-index-id \
        --query-text "How does Formula One work?" \
        --region aws-region
```

Où :

- *kendra-index-id*est votre sauvegardekendra-index-id,
- *aws-region*est votre AWS région.

Windows

```
aws kendra query ^
        --index-id kendra-index-id ^
        --query-text "How does Formula One work?" ^
        --region aws-region
```

Où :

- *kendra-index-id*est votre sauvegardekendra-index-id,
- *aws-region*est votre AWS région.

AWS CLI Affiche les résultats de votre requête.

3. Pour exécuter un exemple de recherche par mot clé, utilisez la commande query :

Linux

```
aws kendra query \
        --index-id kendra-index-id \
        --query-text "Formula One" \
        --region aws-region
```

Où :

- *kendra-index-id*est votre sauvegardekendra-index-id,
- *aws-region*est votre AWS région.

macOS

```
aws kendra query \
        --index-id kendra-index-id \
        --query-text "Formula One" \
        --region aws-region
```

Où :

- *kendra-index-id*est votre sauvegardekendra-index-id,
- *aws-region*est votre AWS région.

Windows

```
aws kendra query ^
        --index-id kendra-index-id ^
        --query-text "Formula One" ^
        --region aws-region
```

Où :

- *kendra-index-id*est votre sauvegardekendra-index-id,
- *aws-region*est votre AWS région.

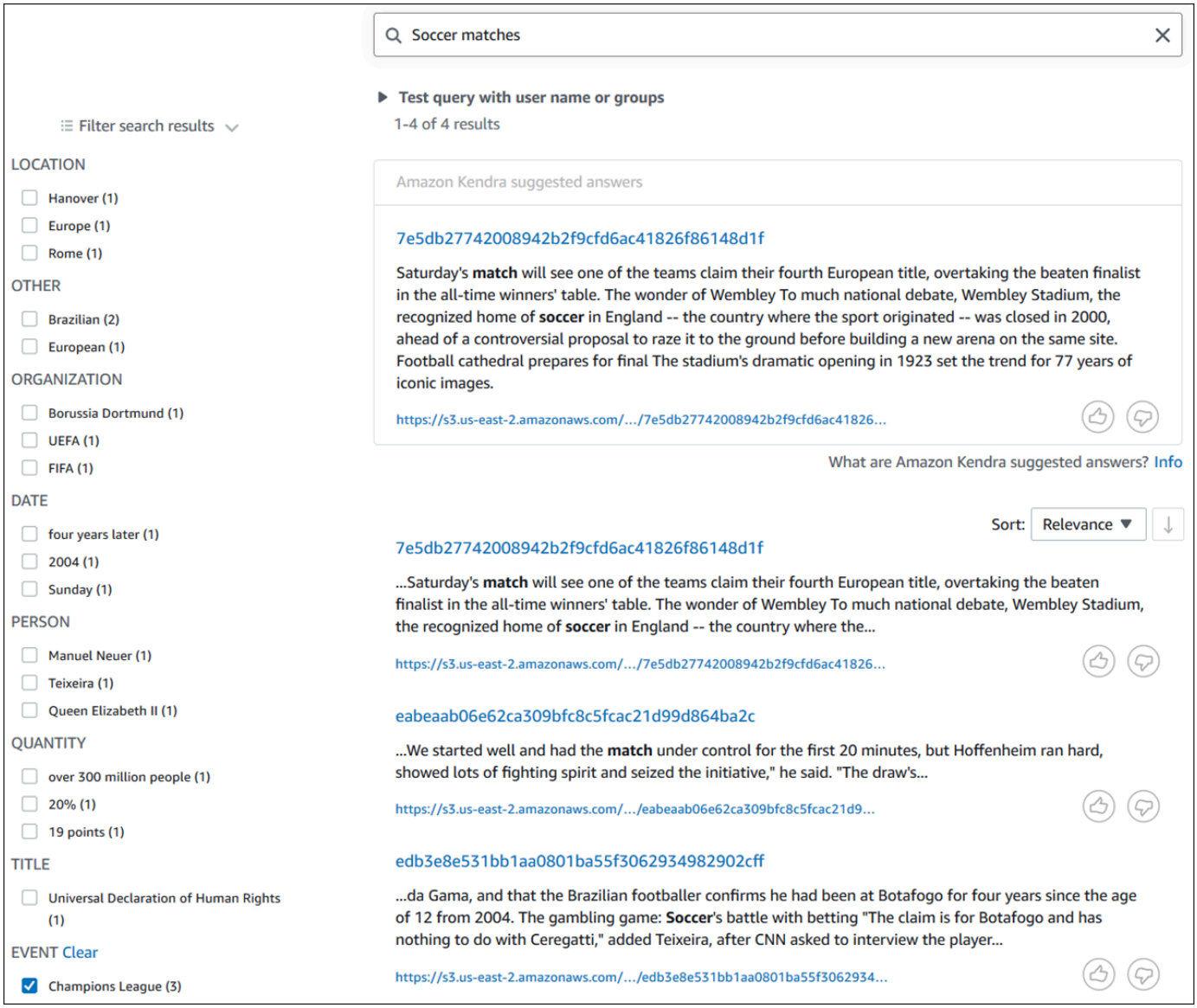
AWS CLI Affiche les réponses renvoyées à votre requête.

Filtrer les résultats de recherche

Vous pouvez filtrer et trier les résultats de recherche à l'aide d'attributs de document personnalisés dans la console Amazon Kendra. Pour plus d'informations sur la façon dont Amazon Kendra traite les requêtes, consultez la section Filtrage des requêtes.

Pour filtrer les résultats de recherche (Console)

1. Ouvrez la console Amazon Kendra à l'adresse. https://console.aws.amazon.com/kendra/
2. Dans la liste des index, cliquez sur`kendra-index`.
3. Dans le menu de navigation de gauche, choisissez l'option de recherche dans votre index.
4. Dans le champ de recherche, entrez **`Soccer matches`** sous forme de requête et appuyez sur Entrée.
5. Dans le menu de navigation de gauche, choisissez Filtrer les résultats de recherche pour afficher une liste de facettes que vous pouvez utiliser pour filtrer votre recherche.
6. Cochez la case « Ligue des champions » sous le sous-titre ÉVÉNEMENT, pour que les résultats de votre recherche soient filtrés uniquement en fonction des résultats contenant « Ligue des champions ».



Pour filtrer les résultats de votre recherche (AWS CLI)

1. Pour voir les entités d'un type spécifique (par exemple`EVENT`) disponibles pour une recherche, utilisez la commande query :

   Linux

```
aws kendra query \
        --index-id kendra-index-id \
        --query-text "Soccer matches" \
        --facets '[{"DocumentAttributeKey":"EVENT"}]' \
        --region aws-region
```

Où :

- *kendra-index-id*est votre sauvegardekendra-index-id,
- *aws-region*est votre AWS région.

macOS

```
aws kendra query \
        --index-id kendra-index-id \
        --query-text "Soccer matches" \
        --facets '[{"DocumentAttributeKey":"EVENT"}]' \
        --region aws-region
```

Où :

- *kendra-index-id*est votre sauvegardekendra-index-id,
- *aws-region*est votre AWS région.

Windows

```
aws kendra query ^
        --index-id kendra-index-id ^
        --query-text "Soccer matches" ^
        --facets '[{"DocumentAttributeKey":"EVENT"}]' ^
        --region aws-region
```

Où :

- *kendra-index-id*est votre sauvegardekendra-index-id,
- *aws-region*est votre AWS région.

AWS CLI Affiche les résultats de la recherche. Pour obtenir une liste des facettes par typeEVENT, accédez à la section « FacetResults » de la AWS CLI sortie pour voir une liste des facettes filtrables avec leur nombre. Par exemple, l'une des facettes est la « Ligue des champions ».

> **Note**
>
> Au lieu de celaEVENT, vous pouvez choisir l'un des champs d'index que vous avez créés the section called “Création d'un index Amazon Kendra” pour la DocumentAttributeKey valeur.

2. Pour exécuter la même recherche mais filtrer uniquement en fonction des résultats contenant « Champions League », utilisez la commande query :

   Linux

   ```
   aws kendra query \
           --index-id kendra-index-id \
           --query-text "Soccer matches" \
           --attribute-filter '{"ContainsAny":{"Key":"EVENT","Value":
   {"StringListValue":["Champions League"]}}}' \
           --region aws-region
   ```

   Où :

   - *kendra-index-id*est votre sauvegardekendra-index-id,
   - *aws-region*est votre AWS région.

   macOS

   ```
   aws kendra query \
           --index-id kendra-index-id \
           --query-text "Soccer matches" \
           --attribute-filter '{"ContainsAny":{"Key":"EVENT","Value":
   {"StringListValue":["Champions League"]}}}' \
           --region aws-region
   ```

   Où :

   - *kendra-index-id*est votre sauvegardekendra-index-id,
   - *aws-region*est votre AWS région.

Windows

```
aws kendra query ^
        --index-id kendra-index-id ^
        --query-text "Soccer matches" ^
        --attribute-filter '{"ContainsAny":{"Key":"EVENT","Value":
{"StringListValue":["Champions League"]}}}' ^
        --region aws-region
```

Où :

- *kendra-index-id*est votre sauvegardekendra-index-id,
- *aws-region*est votre AWS région.

AWS CLI Affiche les résultats de recherche filtrés.

Étape 6 : Nettoyage

Nettoyage de vos fichiers

Pour arrêter de débiter votre AWS compte une fois que vous aurez terminé ce didacticiel, vous pouvez suivre les étapes suivantes :

1. Supprimer votre compartiment Amazon S3

   Pour plus d'informations sur la suppression d'un bucket, consultez la section Suppression d'un bucket.

2. Supprimer votre index Amazon Kendra

   Pour plus d'informations sur la suppression d'un index Amazon Kendra, consultez Supprimer un index.

3. Supprimer **converter.py**

   - Pour la console : Accédez à AWS CloudShellet assurez-vous que la région correspond à votre AWS région. Une fois le shell bash chargé, tapez la commande suivante dans l'environnement et appuyez sur Entrée.

```
rm converter.py
```

- Pour AWS CLI : Exécutez la commande suivante dans une fenêtre de terminal.

  Linux

  ```
  rm file/converter.py
  ```

  Où :

  - *file/*est le chemin du fichier vers votre `converter.py` appareil local.

  macOS

  ```
  rm file/converter.py
  ```

  Où :

  - *file/*est le chemin du fichier vers votre `converter.py` appareil local.

  Windows

  ```
  rm file/converter.py
  ```

  Où :

  - *file/*est le chemin du fichier vers votre `converter.py` appareil local.

En savoir plus

Pour en savoir plus sur l'intégration d'Amazon Kendra dans votre flux de travail, vous pouvez consulter les articles de blog suivants :

- Balisage des métadonnées du contenu pour une recherche améliorée
- Créez une solution de recherche intelligente avec enrichissement automatique du contenu

Pour en savoir plus sur Amazon Comprehend, vous pouvez consulter le guide du développeur Amazon Comprehend.

Surveillance et journalisation pour Amazon Kendra

Rubriques

- Surveillance de votre index (console)
- Enregistrement des appels d'API Amazon Kendra avec des journaux AWS CloudTrail
- Enregistrement des appels de l'API Amazon Kendra Intelligent Ranking avec des journaux AWS CloudTrail
- Surveiller Amazon Kendra avec Amazon CloudWatch
- Surveillance d'Amazon Kendra avec Amazon Logs CloudWatch

Surveillance de votre index (console)

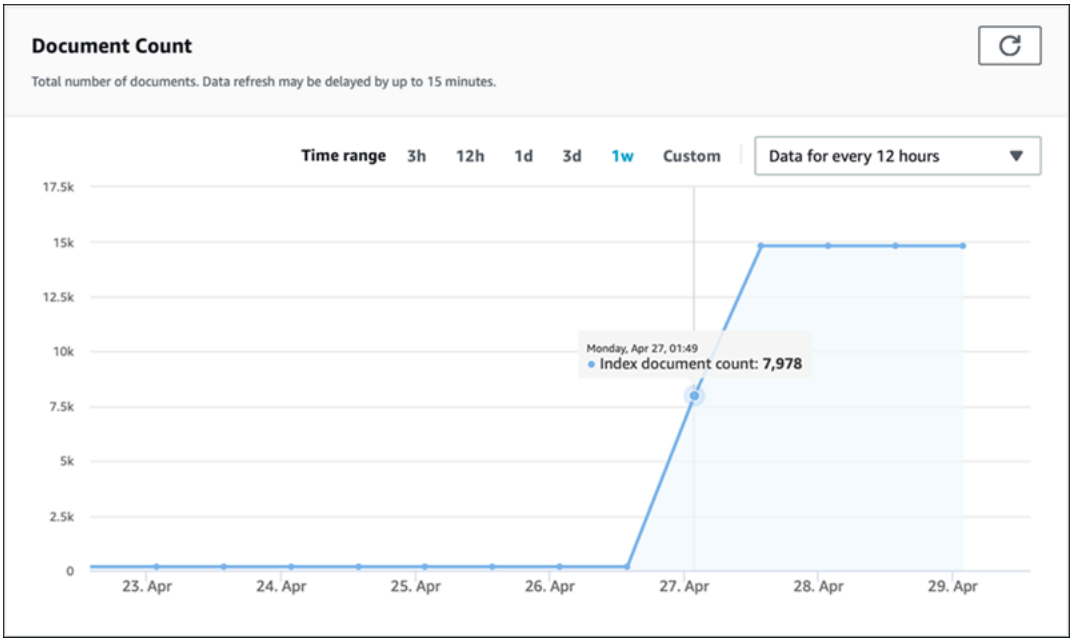
Utilisez la console Amazon Kendra pour surveiller l'état des index et des sources de données. Vous pouvez utiliser ces informations pour suivre la taille et les exigences de stockage de votre index et pour suivre la progression et le succès de la synchronisation entre votre index et les sources de données.

Pour consulter les statistiques de l'indice (console)

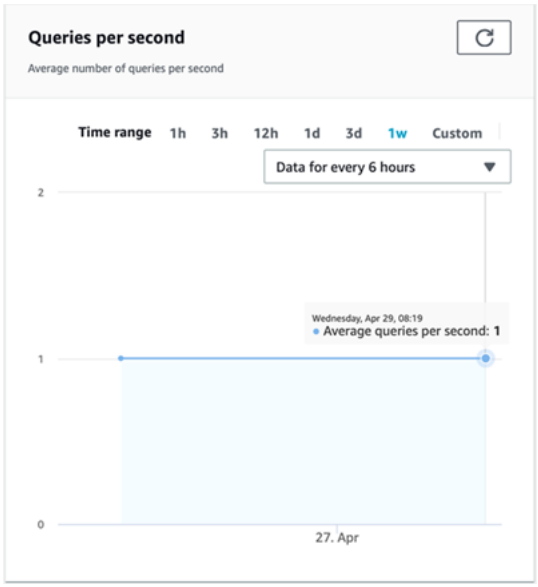
1. Connectez-vous à la console Amazon Kendra AWS Management Console et ouvrez-la chez https://console.aws.amazon.com/kendra/ vous.
2. Dans la liste des index, choisissez l'index à afficher.
3. Faites défiler l'écran pour voir les indicateurs de l'indice.

Vous pouvez consulter les statistiques suivantes concernant votre indice.

- Nombre de documents : nombre total de documents indexés. Cela inclut tous les documents provenant de toutes les sources de données. Utilisez cette métrique pour déterminer si vous devez acheter plus ou moins d'unités de stockage pour votre index.



- Requêtes par seconde : nombre de requêtes d'index demandées chaque seconde. Utilisez cette métrique pour déterminer si vous devez acheter plus ou moins d'unités de requête pour votre index.



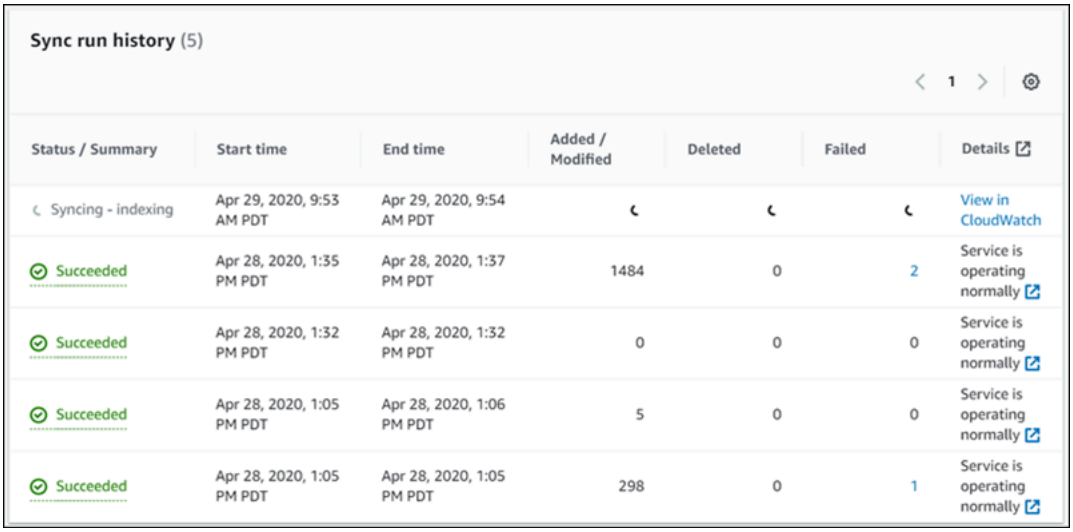
Pour suivre la progression et le succès de la synchronisation entre votre index et une source de données, utilisez la console Amazon Kendra. Utilisez ces informations pour déterminer l'état de santé de votre source de données.

Pour afficher les métriques de synchronisation (console)

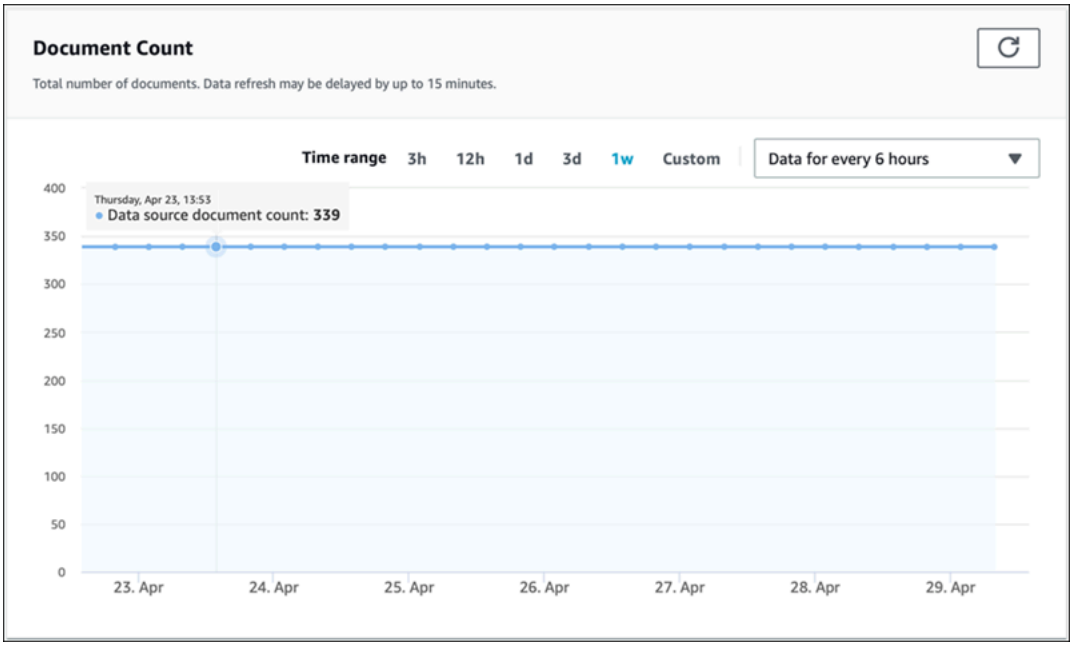
1. Connectez-vous à la console Amazon Kendra AWS Management Console et ouvrez-la chez https://console.aws.amazon.com/kendra/ vous.
2. Dans la liste des index, choisissez l'index pour lequel vous souhaitez afficher les métriques de synchronisation.
3. Dans le menu de gauche, sélectionnez Sources de données.
4. Dans la liste des sources de données, choisissez la source de données à afficher.
5. Faites défiler l'écran pour voir les statistiques relatives à l'exécution de la synchronisation.

Vous pouvez consulter les informations suivantes.

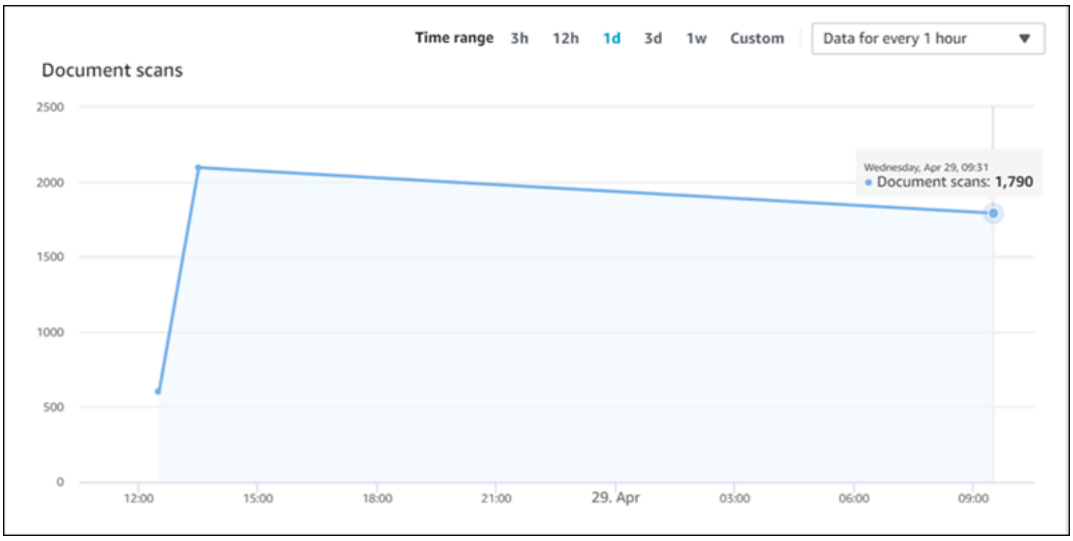
- Historique des opérations de synchronisation : statistiques relatives à l'exécution de la synchronisation, notamment l'heure de début et de fin, le nombre de documents ajoutés, supprimés et ayant échoué. Si la synchronisation échoue, il existe un lien vers les CloudWatch journaux contenant des informations supplémentaires. Cliquez sur l'icône des paramètres en haut à gauche pour modifier les colonnes affichées dans l'historique. Utilisez ces informations pour déterminer l'état général de votre source de données.



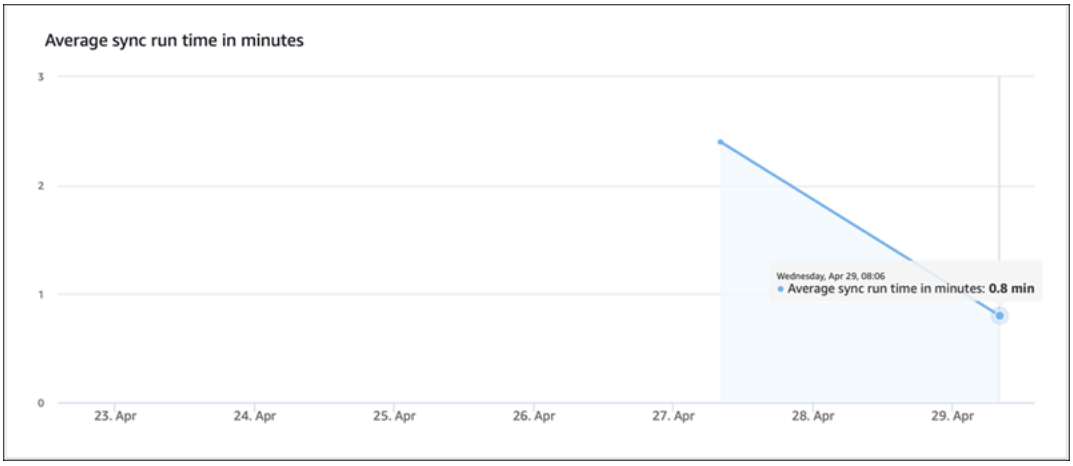
- Nombre de documents : nombre total de documents indexés à partir de cette source de données. Il s'agit du total de tous les documents ajoutés à la source de données moins le total de tous les documents supprimés de la source de données. Utilisez ces informations pour déterminer le nombre de documents issus de cette source de données inclus dans l'index.



- Numérisation de documents : nombre total de documents numérisés lors de la synchronisation. Cela inclut tous les documents de la source de données, y compris ceux ajoutés, mis à jour, supprimés ou inchangés. Utilisez ces informations pour déterminer si Amazon Kendra scanne tous les documents de la source de données. Le nombre de documents numérisés influe sur le montant facturé pour le service.



- Durée moyenne d'exécution de synchronisation en minutes : durée moyenne nécessaire à la fin d'une opération de synchronisation. Le temps nécessaire à la synchronisation d'une source de données influe sur le montant facturé pour le service.



Enregistrement des appels d'API Amazon Kendra avec des journaux AWS CloudTrail

Amazon Kendra est intégré à AWS CloudTrail un service qui fournit un enregistrement des actions effectuées par un utilisateur, un rôle ou un AWS service dans Amazon Kendra. CloudTrail capture tous les appels d'API d'Amazon Kendra sous forme d'événements, y compris les appels depuis la console Amazon Kendra et les appels de code vers Amazon Kendra. APIs Si vous créez un suivi, vous pouvez activer la diffusion continue d' CloudTrail événements vers un compartiment Amazon S3, y compris des événements pour Amazon Kendra. Si vous ne configurez pas de suivi, vous pouvez toujours consulter les événements les plus récents dans la CloudTrail console dans Historique des événements. À l'aide des informations collectées par CloudTrail, vous pouvez déterminer la demande qui a été faite à Amazon Kendra, l'adresse IP à partir de laquelle la demande a été faite, l'auteur de la demande, la date à laquelle elle a été faite, ainsi que des informations supplémentaires.

Pour en savoir plus CloudTrail, notamment comment le configurer et l'activer, consultez le guide de AWS CloudTrail l'utilisateur.

Informations sur Amazon Kendra dans CloudTrail

CloudTrail est activé sur votre AWS compte lorsque vous le créez. Lorsqu'une activité a lieu sur Amazon Kendra, cette activité est enregistrée dans un CloudTrail événement avec d'autres événements de AWS service dans l' CloudTrail historique des événements. Vous pouvez consulter, rechercher et télécharger les événements récents dans votre AWS compte. Pour plus d'informations, consultez Affichage des événements avec l'historique des événements CloudTrail.

Pour un enregistrement continu des événements de votre AWS compte, y compris des événements liés à Amazon Kendra, créez un historique. Un trail est une configuration qui permet de CloudTrail transmettre des événements sous forme de fichiers journaux à un compartiment S3 spécifié. Par défaut, lorsque vous créez un journal de suivi dans la console, il s'applique à toutes les régions AWS . Le journal enregistre les événements de toutes les régions de la AWS partition et transmet les fichiers journaux au compartiment S3 que vous spécifiez. En outre, vous pouvez configurer d'autres AWS services pour analyser plus en détail les données d'événements collectées dans les CloudTrail journaux et agir en conséquence. Pour plus d'informations, consultez :

- Présentation de la création d'un journal d'activité
- CloudTrail Services et intégrations pris en charge
- Configuration des notifications Amazon SNS pour CloudTrail
- Réception de fichiers CloudTrail journaux de plusieurs régions et réception de fichiers CloudTrail journaux de plusieurs comptes

CloudTrail enregistre toutes les actions Amazon Kendra, qui sont documentées dans la référence d'API. Par exemple, les appels aux `Query` opérations `CreateIndexCreateDataSource`, et génèrent des entrées dans les fichiers CloudTrail journaux.

Chaque événement ou entrée de journal contient des informations sur la personne ayant initié la demande. Pour plus d'informations, consultez la section Élément userIdentity CloudTrail.

Exemple : entrées du fichier journal Amazon Kendra

Un suivi est une configuration qui permet de transmettre des événements sous forme de fichiers journaux à un compartiment S3 spécifié. CloudTrail les fichiers journaux contiennent une ou plusieurs entrées de journal. Un événement représente une demande unique provenant de n'importe quelle source et inclut des informations sur l'action demandée, la date et l'heure de l'action, les paramètres de la demande, etc. CloudTrail les fichiers journaux ne constituent pas une trace ordonnée des appels d'API publics, ils n'apparaissent donc pas dans un ordre spécifique.

Les appels à l'`Query`opération créent l'entrée suivante.

```
        {
            "eventVersion": "1.05",
            "userIdentity": {
                "type": "AssumedRole | FederatedUser | IAMUser | Root | SAMLUser |
WebIdentityUser",
```

```
                "principalId": "principal ID",
                "arn": "ARN",
                "accountId": "account ID",
                "accessKeyId": "access key ID",
                "sessionContext": {
                    "sessionIssuer": {
                        "type": "Role",
                        "principalId": "principal Id",
                        "arn": "ARN",
                        "accountId": "account ID",
                        "userName": "user name"
                    },
                    "webIdFederationData": {

                    },
                    "attributes": {
                        "mfaAuthenticated": false,
                        "creationDate": "timestamp"
                    }
                }
            },
            "eventTime": "timestamp",
            "eventSource": "kendra.amazonaws.com",
            "eventName": "Query",
            "awsRegion": "region",
            "sourceIPAddress": "source IP address",
            "userAgent": "user agent",
            "requestParameters": {
                "indexId": "index ID"
            },
            "responseElements": null,
            "requestID": "request ID",
            "eventID": "event ID",
            "eventType": "AwsApiCall",
            "recipientAccountId": "account ID"
        },
```

Enregistrement des appels de l'API Amazon Kendra Intelligent Ranking avec des journaux AWS CloudTrail

Amazon Kendra Intelligent Ranking est intégré à AWS CloudTrail un service qui fournit un enregistrement des actions effectuées par un utilisateur, un rôle ou un AWS service dans Amazon

Kendra Intelligent Ranking. CloudTrail capture tous les appels d'API depuis Amazon Kendra intelligent Ranking sous forme d'événements, y compris les appels de code vers Amazon Kendra Intelligent Ranking. APIs Si vous créez un suivi, vous pouvez activer la diffusion continue d' CloudTrail événements vers un compartiment Amazon S3, y compris des événements pour Amazon Kendra Intelligent Ranking. Si vous ne configurez pas de suivi, vous pouvez toujours consulter les événements les plus récents dans la CloudTrail console dans Historique des événements. À l'aide des informations collectées par CloudTrail, vous pouvez déterminer la demande envoyée à Amazon Kendra Intelligent Ranking, l'adresse IP à partir de laquelle la demande a été faite, l'auteur de la demande, la date à laquelle elle a été faite et des informations supplémentaires.

Pour en savoir plus CloudTrail, notamment comment le configurer et l'activer, consultez le guide de AWS CloudTrail l'utilisateur.

Informations de classement intelligent d'Amazon Kendra dans CloudTrail

CloudTrail est activé sur votre AWS compte lorsque vous le créez. Lorsqu'une activité se produit dans Amazon Kendra Intelligent Ranking, cette activité est enregistrée dans un CloudTrail événement avec d'autres événements de AWS service dans l' CloudTrail historique des événements. Vous pouvez afficher, rechercher et télécharger les événements récents dans votre compte AWS . Pour plus d'informations, consultez la section Affichage des événements à l'aide de l'historique des CloudTrail événements.

Pour un enregistrement continu des événements enregistrés sur votre AWS compte, y compris les événements liés au classement intelligent Amazon Kendra, créez un historique. Un trail est une configuration qui permet de CloudTrail transmettre des événements sous forme de fichiers journaux à un compartiment S3 spécifié. Par défaut, lorsque vous créez un parcours dans la console, celui-ci s'applique à toutes les AWS régions. Le journal enregistre les événements de toutes les régions de la AWS partition et transmet les fichiers journaux au compartiment S3 que vous spécifiez. En outre, vous pouvez configurer d'autres AWS services pour analyser plus en détail les données d'événements collectées dans les CloudTrail journaux et agir en conséquence. Pour plus d’informations, consultez :

- Présentation de la création d’un journal d’activité
- CloudTrail Services et intégrations pris en charge
- Configuration des notifications Amazon SNS pour CloudTrail
- Réception de fichiers CloudTrail journaux de plusieurs régions et réception de fichiers CloudTrail journaux de plusieurs comptes

CloudTrail enregistre toutes les actions d'Amazon Kendra Intelligent Ranking, qui sont documentées dans la référence d'API. Par exemple, les appels aux entrées de `CreateRescoreExecutionPlan` génération dans les fichiers CloudTrail journaux.

Chaque événement ou entrée de journal contient des informations sur la personne ayant initié la demande. Pour plus d'informations, consultez la section Élément userIdentity CloudTrail.

Exemple : entrées du fichier journal Amazon Kendra Intelligent Ranking

Un suivi est une configuration qui permet de transmettre des événements sous forme de fichiers journaux à un compartiment S3 spécifié. CloudTrail les fichiers journaux contiennent une ou plusieurs entrées de journal. Un événement représente une demande unique provenant de n'importe quelle source et inclut des informations sur l'action demandée, la date et l'heure de l'action, les paramètres de la demande, etc. CloudTrail les fichiers journaux ne constituent pas une trace ordonnée des appels d'API publics, ils n'apparaissent donc pas dans un ordre spécifique.

Les appels à l'`CreateRescoreExecutionPlan`opération créent l'entrée suivante.

```
{
            "eventVersion": "1.08",
            "userIdentity": {
                "type": "AssumedRole",
                "principalId": "principal ID",
                "arn": "ARN",
                "accountId": "account ID",
                "accessKeyId": "access key ID",
                "sessionContext": {
                    "sessionIssuer": {
                        "type": "Role",
                        "principalId": "principal ID",
                        "arn": "ARN",
                        "accountId": "account ID",
                        "userName": "user name"
                    },
                    "webIdFederationData": {},
                    "attributes": {
                        "creationDate": "yyyy-mm-ddThh:mm:ssZ",
                        "mfaAuthenticated": "false"
                    }
                }
            },
            "eventTime": "yyyy-mm-ddThh:mm:ssZ",
```

```
            "eventSource": "kendra-ranking.amazonaws.com",
            "eventName": "CreateRescoreExecutionPlan",
            "awsRegion": "region",
            "sourceIPAddress": "source IP address",
            "userAgent": "user agent",
            "requestParameters": {
                "name": "name",
                "description": "description",
                "clientToken": "client token"
            },
            "responseElements": {
                "id": "rescore execution plan ID",
                "arn": "rescore execution plan ARN"
            },
            "requestID": "request ID",
            "eventID": "event ID",
            "readOnly": false,
            "eventType": "AwsApiCall",
            "managementEvent": true,
            "recipientAccountId": "account ID",
            "eventCategory": "Management",
            "tlsDetails": {
                "tlsVersion": "TLS version",
                "cipherSuite": "cipher suite",
                "clientProvidedHostHeader": "kendra-ranking.[region].api.aws"
            }
        }
```

Surveiller Amazon Kendra avec Amazon CloudWatch

Pour suivre l'état de vos index, utilisez Amazon CloudWatch. Avec CloudWatch, vous pouvez obtenir des métriques pour la synchronisation des documents pour votre index. Vous pouvez également configurer des CloudWatch alarmes pour être averties lorsqu'une ou plusieurs mesures dépassent un seuil que vous définissez. Par exemple, vous pouvez contrôler le nombre de documents soumis pour être indexés ou le nombre de documents qui n'ont pas pu être indexés.

Vous devez disposer des CloudWatch autorisations appropriées pour surveiller Amazon Kendra avec. CloudWatch Pour plus d'informations, consultez Authentification et contrôle d'accès pour Amazon CloudWatch dans le guide de CloudWatch l'utilisateur Amazon.

Consulter les statistiques d'Amazon Kendra

Consultez les statistiques d'Amazon Kendra à l'aide de la CloudWatch console.

Pour consulter les métriques (CloudWatch console)

1. Connectez-vous à la CloudWatch console AWS Management Console et ouvrez-la à l'adresse https://console.aws.amazon.com/cloudwatch/.
2. Choisissez Metrics, choisissez All Metrics, puis Kendra.
3. Choisissez la dimension, le nom de la métrique, puis Ajouter au graphique.
4. Choisissez une valeur pour la plage de dates. Le décompte de la métrique pour la plage de dates sélectionnée est affiché dans le graphique.

Création d'une alarme

Une CloudWatch alarme surveille une seule métrique sur une période spécifiée et exécute une ou plusieurs actions : envoyer une notification à un responsable d'Amazon Simple Notification Service (Amazon SNS) ou à une politique Auto Scaling. Les actions ou actions sont basées sur la valeur de la métrique par rapport à un seuil donné sur un certain nombre de périodes que vous spécifiez. CloudWatch peut également vous envoyer un message Amazon SNS lorsque l'alarme change d'état.

CloudWatch les alarmes appellent des actions uniquement lorsque l'état change et persiste pendant la période que vous spécifiez.

Pour définir une alarme

1. Connectez-vous à la CloudWatch console AWS Management Console et ouvrez-la à l'adresse https://console.aws.amazon.com/cloudwatch/.
2. Choisissez Alarmes, puis choisissez Créer une alarme.
3. Sélectionnez une métrique. Choisissez une métrique Kendra pour votre index et votre source de données. Définissez également l'heure sous forme de nombre d'heures, de jours, de semaines défini ou sur mesure.
4. Choisissez votre statistique. Par exemple, Average. Choisissez également la période de déclenchement de votre alarme sous forme de nombre défini de minutes, d'heures, par jour ou sur mesure.
5. Choisissez votre seuil pour déclencher l'alarme, qu'il s'agisse d'une valeur statique ou d'une bande, ainsi que la condition à respecter pour le seuil.

6. Choisissez l'état d'alarme pour le déclencheur, si la métrique doit dépasser le seuil que vous avez défini, ou un autre état. Sélectionnez à qui/à quel e-mail envoyer la notification d'alarme.
7. Si vous êtes satisfait de l'alarme, choisissez Créer une alarme.

> **Note**
>
> Vous devez donner un nom à votre CloudWatch alarme.

CloudWatch Mesures pour les tâches de synchronisation d'index

Le tableau suivant décrit les métriques Amazon Kendra pour les tâches de synchronisation des sources de données.

Si vous utilisez l'API ou la CLI, vous devez spécifier « AWS/Kendra » en plus de celui `MetricName` de votre choix lorsque vous utilisez l'API. `Namespace` GetMetricStatistics

Métrique	Description
`DocumentsCrawled`	Le nombre de documents que la tâche de synchronisation a scannés ou découverts pendant l'exécution. Dimensions : • IndexId • DataSourceId Unité : nombre
`DocumentsSubmittedForIndexing`	Nombre de documents que la tâche de synchronisation a soumis à l'index. Dimensions : • IndexId • DataSourceId

Métrique	Description
	Unité : nombre
DocumentsSubmittedForIndexingFailed	Nombre de documents dont l'indexation a échoué. Consultez le contenu du CloudWatch journal de la tâche de synchronisation pour plus de détails. Dimensions : • IndexId • DataSourceId Unité : nombre
DocumentsSubmittedForDeletion	Nombre de documents que la tâche de synchronisation a demandé de supprimer de l'index. Dimensions : • IndexId • DataSourceId Unité : nombre
DocumentsSubmittedForDeletionFailed	Le nombre de documents qui n'ont pas pu être supprimés. Consultez le contenu du CloudWatch journal de la tâche de synchronisation pour plus de détails. Dimensions : • IndexId • DataSourceId Unité : nombre

Mesures pour les sources de données Amazon Kendra

Le tableau suivant décrit les métriques Amazon Kendra pour les tâches de synchronisation des sources de données. Les métriques marquées d'un astérisque (*) sont utilisées uniquement pour les sources de données Amazon S3.

Si vous utilisez l'API ou la CLI, vous devez spécifier « AWS/Kendra » en plus de celui `MetricName` de votre choix lorsque vous utilisez l'API. `Namespace` GetMetricStatistics

Métrique	Description
`DocumentsSkippedNoChange *`	Le nombre de documents examinés et jugés inchangés, c'est-à-dire qu'ils n'ont pas été soumis à des fins d'indexation. Dimensions : • IndexId • DataSourceId Unité : nombre
`DocumentsSkippedInvalidMetadata *`	Le nombre de documents ignorés en raison d'un problème lié au fichier de métadonnées associé. Consultez le contenu du CloudWatch journal de l'exécution de synchronisation pour plus de détails. Dimensions : • IndexId • DataSourceId Unité : nombre
`DocumentsCrawled`	Le nombre de fichiers de documents examinés. Dimensions :

Métrique	Description
	• IndexId • DataSourceId Unité : nombre
DocumentsSubmittedForDeletion	Le nombre de documents examinés qui ont été supprimés de la source de données et soumis pour suppression. Dimensions : • IndexId • DataSourceId Unité : nombre
DocumentsSubmittedForDeletionFailed	Nombre de documents dont la suppression d'une source de données a échoué. Dimensions : • IndexId • DataSourceId Unité : nombre
DocumentsSubmittedForIndexing	Le nombre de documents examinés et soumis pour indexation. Dimensions : • IndexId • DataSourceId Unité : nombre

Métrique	Description
`DocumentsSubmittedForIndexingFailed`	Le nombre de documents soumis pour indexation qui n'ont pas pu être indexés. Dimensions : • IndexId • DataSourceId Unité : nombre

Mesures pour les documents indexés

Le tableau suivant décrit les statistiques Amazon Kendra pour les documents indexés. Pour les documents indexés à l'aide de l'BatchPutDocumentopération, seule la `IndexId` dimension est prise en charge.

Si vous utilisez l'API ou la CLI, vous devez spécifier « AWS/Kendra » en plus de celui `MetricName` de votre choix lorsque vous utilisez l'API. `Namespace` GetMetricStatistics

Métrique	Description
`DocumentsIndexed`	Le nombre de documents indexés. Dimensions : • IndexId • DataSourceId Unité : nombre
`DocumentsFailedToIndex`	Nombre de documents qui n'ont pas pu être indexés. Consultez le contenu du CloudWatch journal pour plus de détails. Dimensions :

Métrique	Description
	• IndexId • DataSourceId Unité : nombre
`IndexQueryCount`	Nombre de requêtes d'index par minute. Dimensions : • IndexId Unité : nombre

Surveillance d'Amazon Kendra avec Amazon Logs CloudWatch

Amazon Kendra utilise Amazon CloudWatch Logs pour vous donner un aperçu du fonctionnement de vos sources de données. Amazon Kendra enregistre les détails du processus des documents au fur et à mesure de leur indexation. Il enregistre les erreurs de votre source de données qui se produisent lors de l'indexation de vos documents. Vous utilisez CloudWatch les journaux pour surveiller, stocker et accéder aux fichiers journaux.

CloudWatch Les journaux stockent les événements des journaux dans un flux de journaux faisant partie d'un groupe de journaux. Amazon Kendra utilise ces fonctionnalités comme suit :

- Groupes de journaux : Amazon Kendra stocke tous vos flux de journaux dans un seul groupe de journaux pour chaque index. Amazon Kendra crée le groupe de journaux lors de la création de l'index. L'identifiant du groupe de logs commence toujours par « aws/kendra/ ».
- Flux de journaux : Amazon Kendra crée un nouveau flux de journal de source de données dans le groupe de journaux pour chaque tâche de synchronisation d'index que vous exécutez. Il crée également un nouveau flux de journal de documents lorsqu'un flux atteint environ 500 entrées.
- Entrées de journal : Amazon Kendra crée une entrée de journal dans le flux de journal lors de l'indexation des documents. Chaque entrée fournit des informations sur le traitement du document ou sur les éventuelles erreurs rencontrées.

Pour plus d'informations sur l'utilisation CloudWatch des journaux, consultez la section Qu'est-ce qu'Amazon Cloud Watch Logs dans le guide de l'utilisateur d'Amazon Cloud Watch Logs.

Amazon Kendra crée deux types de flux de journaux :

- Flux de journaux des sources de données
- Flux de journaux de documents

Flux de journaux des sources de données

Les flux de journaux des sources de données publient des entrées relatives à vos tâches de synchronisation d'index. Chaque tâche de synchronisation crée un nouveau flux de journal qu'elle utilise pour publier des entrées. Le nom du flux de log est le suivant :

```
data source id/YYYY-MM-DD-HH/data source sync job ID
```

Un nouveau flux de journal est créé pour chaque tâche de synchronisation exécutée.

Il existe trois types de messages de journal publiés dans le flux de journal d'une source de données :

- Message de journal pour un document qui n'a pas pu être envoyé pour indexation. Voici un exemple de ce message pour un document dans une source de données S3 :

```
{
    "DocumentId": "document ID",
    "S3Path": "s3://bucket/prefix/object",
    "Message": "Failed to ingest document via BatchPutDocument.",
    "ErrorCode": "InvalidRequest",
    "ErrorMessage": "No document metadata configuration found for document attribute
 key  city."
}
```

- Message de journal concernant un document qui n'a pas pu être envoyé pour suppression. Voici un exemple de ce message :

```
{
    "DocumentId": "document ID",
    "Message": "Failed to delete document via BatchDeleteDocument.",
    "ErrorCode": "InvalidRequest",
    "ErrorMessage": "Document can't be deleted because it doesn't exist."
```

```
}
```

- Un message de journal lorsqu'un fichier de métadonnées non valide est détecté pour un document dans un compartiment Amazon S3. Voici un exemple de ce message.

```
{
    "Message": "Found invalid metadata
 file bucket/prefix/filename.extension.metadata.json."
}
```

- Pour les connecteurs de base de données SharePoint et les connecteurs de base de données, Amazon Kendra n'écrit des messages dans le flux de journal que si un document ne peut pas être indexé. Voici un exemple du message d'erreur enregistré par Amazon Kendra.

```
{
    "DocumentID": "document ID",
    "IndexID": "index ID",
    "SourceURI": "",
    "CrawlStatus": "FAILED",
    "ErrorCode": "403",
    "ErrorMessage": "Access Denied",
    "DataSourceErrorCode": "403"
}
```

Flux de journaux de documents

Amazon Kendra enregistre les informations relatives au traitement des documents lors de leur indexation. Il enregistre un ensemble de messages pour les documents stockés dans une source de données Amazon S3. Il enregistre les erreurs uniquement pour les documents stockés dans une source de données Microsoft SharePoint ou une base de données.

Si les documents ont été ajoutés à l'index à l'aide de l'BatchPutDocumentopération, le flux de log est nommé comme suit :

```
YYYY-MM-DD-HH/UUID
```

Si les documents ont été ajoutés à l'index à l'aide d'une source de données, le flux de log est nommé comme suit :

```
dataSourceId/YYYY-MM-DD-HH/UUID
```

Chaque flux de journal contient jusqu'à 500 messages.

Si l'indexation d'un document échoue, le message suivant est envoyé dans le flux du journal :

```
{
    "DocumentId": "document ID",
    "IndexName": "index name",
    "IndexId": "index ID"
    "SourceURI": "source URI"
    "IndexingStatus": "DocumentFailedToIndex",
    "ErrorCode": "400 | 500",
    "ErrorMessage": "message"
}
```

Afficher les statistiques Amazon Kendra pour vos tâches de synchronisation

Vous pouvez consulter un rapport d'historique des opérations de synchronisation au niveau du document dans le cadre CloudWatch de votre tâche de synchronisation des sources de données en sélectionnant Afficher le rapport. Un rapport d'historique des opérations de synchronisation contiendra des détails sur la progression et le statut de chaque document dans le cadre de la tâche de synchronisation. Il indique si un document a réussi, a échoué ou a été ignoré pendant les étapes d'analyse, de synchronisation et d'indexation. Vous trouverez également tous les messages d'erreur relatifs à des documents échoués ou ignorés. Si le rapport n'affiche pas les résultats d'une tâche de synchronisation en cours, il est possible que les journaux ne soient pas encore disponibles. Revenez plus tard au fur et à mesure que des données sont émises dans le rapport lorsque des événements se produisent pendant le processus de synchronisation.

Pour accéder à votre rapport sur l'historique des opérations de synchronisation, procédez comme suit :

1. Ouvrez la console Amazon Kendra à l'adresse. https://console.aws.amazon.com/kendra/
2. Dans le menu de navigation de gauche, sous Gestion des données, choisissez Sources de données, puis choisissez votre source de données.
3. Sur la page récapitulative de votre source de données, faites défiler l'écran vers le bas et sélectionnez l'onglet Historique de synchronisation.
4. Dans l'historique des exécutions de synchronisation, sélectionnez Actions.

5. Dans Actions, sélectionnez Afficher le rapport. Vous serez redirigé vers la CloudWatch console où vous pourrez accéder à votre rapport.

> **Note**
>
> L'historique des opérations de synchronisation enregistre si un document a été correctement indexé lors de l'ingestion, y compris les pièces jointes ACLs et les métadonnées, pour tous les connecteurs pris en charge par Amazon Kendra.

Si vous utilisez le connecteur Amazon S3 :

Outre l'affichage du rapport d'historique des opérations de synchronisation au niveau du document dans CloudWatch, vous pouvez générer des rapports d'historique de synchronisation pour chaque document de votre source de données Amazon S3 et le copier dans un compartiment. Amazon S3 Au cours de ce processus, vos données sont cryptées à l'aide de AWS KMS clés et vous seul pouvez les consulter. Le statut du document signalé peut être l'un des suivants : Echec, Terminé ou Réussite avec des erreurs. Avant de pouvoir générer des rapports d'état de synchronisation pour Amazon S3, vous devez effectuer les opérations suivantes :

- Ajoutez le principal Amazon Kendra de service suivant à votre politique Amazon S3 d'accès

```
{
    “Version”: “2012-10-17",
    “Statement”: [
        {
            “Sid”: “KendraS3Access”,
            “Effect”: “Allow”,
            “Principal”: {
                “Service”: “kendra.amazonaws.com”
            },
            “Action”: “s3:PutObject”,
            “Resource”: “arn:aws:s3:::your-manifest-bucket-name/*”
        }
    ]
}
```

- Créez un Amazon S3 bucket avec des autorisations d'accès pour Amazon Kendra

Si vous utilisez la console, pour générer un rapport d'historique de synchronisation pour Amazon S3, choisissez d'activer l'option Générer des rapports dans la section facultative de synchronisation des rapports d'historique sur la page de détails de la source de données. Entrez ensuite l'emplacement du Amazon S3 compartiment et choisissez parmi les options de configuration disponibles. Les rapports seront générés lors de la prochaine synchronisation une fois que vous aurez activé l'option Générer un rapport.

Si vous supprimez le Amazon S3 compartiment, vous perdrez vos données de journal et devrez en configurer un nouveau pour stocker les nouveaux rapports de synchronisation.

> **Note**
>
> Un rapport d'historique de synchronisation fournit uniquement des informations indiquant si un connecteur Amazon S3 a correctement exploré et ingéré des données.

Sécurité dans Amazon Kendra

La sécurité du cloud AWS est la priorité absolue. En tant que AWS client, vous bénéficiez d'un centre de données et d'une architecture réseau conçus pour répondre aux exigences des entreprises les plus sensibles en matière de sécurité.

La sécurité est une responsabilité partagée entre vous AWS et vous. Le modèle de responsabilité partagée décrit cela comme la sécurité du cloud et la sécurité dans le cloud :

- Sécurité du cloud :AWS est chargée de protéger l'infrastructure qui exécute les AWS services dans le AWS cloud. AWS vous fournit également des services que vous pouvez utiliser en toute sécurité. Des auditeurs tiers testent et vérifient régulièrement l'efficacité de notre sécurité dans le cadre des programmes de AWS conformité Programmes de de conformité. Pour en savoir plus sur les programmes de conformité qui s'appliquent à Amazon Kendra, consultez la section AWS Services concernés par programme de conformitéAWS .
- Sécurité dans le cloud : votre responsabilité est déterminée par le AWS service que vous utilisez. Vous êtes également responsable d'autres facteurs, y compris de la sensibilité de vos données, des exigences de votre entreprise, ainsi que de la législation et de la réglementation applicables.

Cette documentation vous aide à comprendre comment appliquer le modèle de responsabilité partagée lors de l'utilisation d'Amazon Kendra. Les rubriques suivantes expliquent comment configurer Amazon Kendra pour répondre à vos objectifs de sécurité et de conformité. Vous apprendrez également à utiliser d'autres AWS services qui vous aident à surveiller et à sécuriser vos ressources Amazon Kendra.

Rubriques

- Protection des données dans Amazon Kendra
- Amazon Kendra Amazon Kendra Classement intelligent et interface VPC ()AWS PrivateLink
- Gestion des identités et des accès pour Amazon Kendra
- Bonnes pratiques de sécurité
- Journalisation et surveillance dans Amazon Kendra
- Validation de conformité pour Amazon Kendra
- La résilience chez Amazon Kendra
- Sécurité de l'infrastructure dans Amazon Kendra
- Analyse de configuration et de vulnérabilité dans AWS Identity and Access Management

Protection des données dans Amazon Kendra

Le modèle de responsabilité AWS partagée s'applique à la protection des données dans Amazon Kendra. Comme décrit dans ce modèle, AWS est chargé de protéger l'infrastructure mondiale qui gère tous les AWS Cloud. La gestion du contrôle de votre contenu hébergé sur cette infrastructure relève de votre responsabilité. Vous êtes également responsable des tâches de configuration et de gestion de la sécurité des Services AWS que vous utilisez. Pour plus d'informations sur la confidentialité des données, consultez Questions fréquentes (FAQ) sur la confidentialité des données. Pour en savoir plus sur la protection des données en Europe, consultez le billet de blog Modèle de responsabilité partagée AWS et RGPD (Règlement général sur la protection des données) sur le Blog de sécuritéAWS .

À des fins de protection des données, nous vous recommandons de protéger les Compte AWS informations d'identification et de configurer les utilisateurs individuels avec AWS IAM Identity Center ou AWS Identity and Access Management (IAM). Ainsi, chaque utilisateur se voit attribuer uniquement les autorisations nécessaires pour exécuter ses tâches. Nous vous recommandons également de sécuriser vos données comme indiqué ci-dessous :

- Utilisez l'authentification multifactorielle (MFA) avec chaque compte.
- SSL/TLS À utiliser pour communiquer avec AWS les ressources. Nous exigeons TLS 1.2 et recommandons TLS 1.3.
- Configurez l'API et la journalisation de l'activité des utilisateurs avec AWS CloudTrail. Pour plus d'informations sur l'utilisation des CloudTrail sentiers pour capturer AWS des activités, consultez la section Utilisation des CloudTrail sentiers dans le guide de AWS CloudTrail l'utilisateur.
- Utilisez des solutions de AWS chiffrement, ainsi que tous les contrôles de sécurité par défaut qu'ils contiennent Services AWS.
- Utilisez des services de sécurité gérés avancés tels qu'Amazon Macie, qui contribuent à la découverte et à la sécurisation des données sensibles stockées dans Amazon S3.
- Si vous avez besoin de modules cryptographiques validés par la norme FIPS 140-3 pour accéder AWS via une interface de ligne de commande ou une API, utilisez un point de terminaison FIPS. Pour plus d'informations sur les points de terminaison FIPS disponibles, consultez Norme FIPS (Federal Information Processing Standard) 140-3.

Nous vous recommandons fortement de ne jamais placer d'informations confidentielles ou sensibles, telles que les adresses e-mail de vos clients, dans des balises ou des champs de texte libre tels que le champ Nom. Cela inclut lorsque vous travaillez avec Amazon Kendra ou une autre entreprise à

Services AWS l'aide de la console, de l'API ou. AWS CLI AWS SDKs Toutes les données que vous entrez dans des balises ou des champs de texte de forme libre utilisés pour les noms peuvent être utilisées à des fins de facturation ou dans les journaux de diagnostic. Si vous fournissez une adresse URL à un serveur externe, nous vous recommandons fortement de ne pas inclure d'informations d'identification dans l'adresse URL permettant de valider votre demande adressée à ce serveur.

Chiffrement au repos

Amazon Kendra chiffre vos données au repos à l'aide de la clé de chiffrement de votre choix. Vous pouvez choisir l'une des méthodes suivantes.

- Une clé AWS KMS AWS détenue par un utilisateur. Si vous ne spécifiez pas de clé de chiffrement, vos données sont chiffrées avec cette clé par défaut.
- Une clé KMS AWS gérée dans votre compte. Cette clé est créée, gérée et utilisée en votre nom par Amazon Kendra. Le nom de la clé est`aws/kendra`.
- Une clé gérée par le client. Vous pouvez fournir l'ARN d'une clé de chiffrement que vous avez créée dans votre compte. Lorsque vous utilisez une clé KMS gérée par le client, vous devez lui attribuer une politique de clé qui autorise Amazon Kendra à l'utiliser. Sélectionnez une clé KMS de chiffrement symétrique gérée par le client, Amazon Kendra ne prend pas en charge les clés KMS asymétriques. Pour de plus amples informations, veuillez consulter Gestion des clés.

Chiffrement en transit

Amazon Kendra utilise le protocole HTTPS pour communiquer avec votre application cliente. Il utilise le protocole HTTPS et AWS des signatures pour communiquer avec d'autres services au nom de votre application. Si vous utilisez un VPC, vous pouvez l'utiliser AWS PrivateLink pour établir une connexion privée entre votre VPC et Amazon Kendra.

Gestion des clés

Amazon Kendra chiffre le contenu de votre index à l'aide de l'un des trois types de clés. Vous pouvez choisir l'une des méthodes suivantes.

- Un AWS KMS AWS appartenant à un propriétaire. Il s'agit de l'option par défaut.
- Une clé KMS AWS gérée. Cette clé est créée dans votre compte et est gérée et utilisée en votre nom par Amazon Kendra.

- Une clé KMS gérée par le client. Vous pouvez créer la clé lorsque vous créez un index ou une source de données Amazon Kendra, ou vous pouvez créer la clé à l'aide de la console. AWS KMS Sélectionnez une clé KMS de chiffrement symétrique gérée par le client. Amazon Kendra ne prend pas en charge les clés KMS asymétriques. Pour plus d'informations, veuillez consulter la rubrique Utilisation des clés symétriques et asymétriques dans le Guide du développeur AWS Key Management Service.

Amazon Kendra Amazon Kendra Classement intelligent et interface VPC ()AWS PrivateLink

Vous pouvez établir une connexion privée entre votre VPC et Amazon Kendra en créant un point de terminaison VPC d'interface. Les points de terminaison de l'interface sont alimentés par AWS PrivateLinkune technologie qui vous permet d'accéder en privé à Amazon APIs Kendra sans passerelle Internet, appareil NAT, connexion VPN ou connexion AWS Direct Connect. Les instances de votre VPC n'ont pas besoin d'adresses IP publiques pour communiquer avec Amazon Kendra. APIs Le trafic entre votre VPC et Amazon Kendra ne quitte pas le réseau Amazon.

Chaque point de terminaison d'interface est représenté par une ou plusieurs interfaces réseau Elastic dans vos sous-réseaux.

Considérations relatives aux points de terminaison VPC Amazon Kendra et Amazon Kendra Intelligent Ranking

Avant de configurer un point de terminaison VPC d'interface pour Amazon Kendra ou Amazon Kendra Intelligent Ranking, assurez-vous de consulter les conditions requises dans le guide de l'utilisateur Amazon VPC.

Amazon Kendra et Amazon Kendra Intelligent Ranking vous permettent d'appeler toutes ses actions d'API depuis votre VPC.

Création d'un point de terminaison VPC d'interface pour Amazon Kendra et Amazon Kendra Intelligent Ranking

Vous pouvez créer un point de terminaison VPC pour le service Amazon Kendra ou Amazon Kendra Intelligent Ranking à l'aide de la console Amazon VPC ou du (). AWS Command Line Interface AWS CLI

Créez un point de terminaison VPC pour Amazon Kendra en utilisant le nom de service suivant :

- com.amazonaws. *region*.kendra

Créez un point de terminaison VPC pour Amazon Kendra Intelligent Ranking en utilisant le nom de service suivant :

- aws.api. *region*classement .kendra

Après avoir créé un point de terminaison VPC, vous pouvez utiliser l'exemple de AWS CLI commande suivant qui utilise le `endpoint-url` paramètre pour spécifier un point de terminaison d'interface pour l'API Amazon Kendra :

```
aws kendra list-indices --endpoint-url https://VPC endpoint
```

*VPC endpoint*est le nom DNS généré lors de la création du point de terminaison de l'interface. Ce nom inclut l'ID du point de terminaison VPC et le nom du service Amazon Kendra, qui inclut la région. Par exemple, `vpce-1234-abcdef.kendra.us-west-2.vpce.amazonaws.com`.

Si vous activez le DNS privé pour le point de terminaison, vous pouvez envoyer des demandes d'API à Amazon Kendra en utilisant son nom DNS par défaut pour la région. Par exemple, `kendra.us-east-1.amazonaws.com`.

Pour plus d'informations, consultez Création d'un point de terminaison d'interface dans le Guide de l'utilisateur Amazon VPC.

Création d'une politique de point de terminaison VPC pour Amazon Kendra et Amazon Kendra Intelligent Ranking

Vous pouvez associer une politique de point de terminaison à votre point de terminaison VPC qui contrôle l'accès à Amazon Kendra ou Amazon Kendra Intelligent Ranking.

La politique relative à Amazon Kendra ou Amazon Kendra Intelligent Ranking précise les informations suivantes :

- L' principal/authorized utilisateur qui peut effectuer des actions.
- Les actions qui peuvent être effectuées.
- Les ressources sur lesquelles les actions peuvent être exécutées.

Exemple : politique de point de terminaison VPC pour les actions Amazon Kendra

Voici un exemple de politique de point de terminaison pour Amazon Kendra. Lorsqu'elle est attachée à un point de terminaison, cette politique donne accès à toutes les actions Amazon Kendra disponibles à tous les principals/authorized utilisateurs sur toutes les ressources.

```
{
   "Statement":[
      {
         "Principal":"*",
         "Effect":"Allow",
         "Action":[
            "kendra:*"
         ],
         "Resource":"*"
      }
   ]
}
```

Exemple : politique de point de terminaison VPC pour les actions de classement intelligent Amazon Kendra

Voici un exemple de politique de point de terminaison pour Amazon Kendra Intelligent Ranking. Lorsqu'elle est associée à un point de terminaison, cette politique donne accès à toutes les actions Amazon Kendra Intelligent Ranking disponibles à tous les principals/authorized utilisateurs sur toutes les ressources.

```
{
   "Statement":[
      {
         "Principal":"*",
         "Effect":"Allow",
         "Action":[
            "kendra-ranking:*"
         ],
         "Resource":"*"
      }
   ]
}
```

Pour plus d'informations, consultez la section Contrôle de l'accès aux points de terminaison VPC à l'aide des politiques relatives aux points de terminaison dans le guide de l'utilisateur Amazon VPC.

Gestion des identités et des accès pour Amazon Kendra

AWS Identity and Access Management (IAM) est un outil Service AWS qui permet à un administrateur de contrôler en toute sécurité l'accès aux AWS ressources. Les administrateurs IAM contrôlent qui peut être authentifié (connecté) et autorisé (autorisé) à utiliser les ressources Amazon Kendra. IAM est un Service AWS outil que vous pouvez utiliser sans frais supplémentaires.

Rubriques

- Public ciblé
- Authentification par des identités
- Gestion des accès à l’aide de politiques
- Comment Amazon Kendra travaille avec IAM
- Exemples de politiques basées sur l'identité d'Amazon Kendra
- AWS politiques gérées pour Amazon Kendra
- Résolution des problèmes liés à l'identité et à l'accès à Amazon Kendra

Public ciblé

La façon dont vous utilisez AWS Identity and Access Management (IAM) varie en fonction du travail que vous effectuez dans Amazon Kendra.

Utilisateur du service : si vous utilisez le service Amazon Kendra pour effectuer votre travail, votre administrateur vous fournit les informations d'identification et les autorisations dont vous avez besoin. Au fur et à mesure que vous utilisez de plus en plus de fonctionnalités d'Amazon Kendra pour effectuer votre travail, vous aurez peut-être besoin d'autorisations supplémentaires. En comprenant bien la gestion des accès, vous saurez demander les autorisations appropriées à votre administrateur. Si vous ne parvenez pas à accéder à une fonctionnalité d'Amazon Kendra, consultez. Résolution des problèmes liés à l'identité et à l'accès à Amazon Kendra

Administrateur du service — Si vous êtes responsable des ressources Amazon Kendra au sein de votre entreprise, vous avez probablement un accès complet à Amazon Kendra. C'est à vous de déterminer les fonctionnalités et les ressources d'Amazon Kendra auxquelles les utilisateurs de vos services doivent accéder. Vous devez ensuite soumettre les demandes à votre administrateur IAM pour modifier les autorisations des utilisateurs de votre service. Consultez les informations sur cette page pour comprendre les concepts de base d’IAM. Pour en savoir plus sur la manière dont votre

entreprise peut utiliser l'IAM avec Amazon Kendra, consultez. Comment Amazon Kendra travaille avec IAM

Administrateur IAM : si vous êtes administrateur IAM, vous souhaiterez peut-être en savoir plus sur la manière dont vous pouvez rédiger des politiques pour gérer l'accès à Amazon Kendra. Pour consulter des exemples de politiques basées sur l'identité d'Amazon Kendra que vous pouvez utiliser dans IAM, consultez. Exemples de politiques basées sur l'identité d'Amazon Kendra

Authentification par des identités

L'authentification est la façon dont vous vous connectez à AWS l'aide de vos informations d'identification. Vous devez être authentifié (connecté à AWS) en tant qu'utilisateur IAM ou en assumant un rôle IAM. Utilisateur racine d'un compte AWS

Vous pouvez vous connecter en AWS tant qu'identité fédérée en utilisant les informations d'identification fournies par le biais d'une source d'identité. AWS IAM Identity Center Les utilisateurs (IAM Identity Center), l'authentification unique de votre entreprise et vos informations d'identification Google ou Facebook sont des exemples d'identités fédérées. Lorsque vous vous connectez avec une identité fédérée, votre administrateur aura précédemment configuré une fédération d'identités avec des rôles IAM. Lorsque vous accédez à AWS l'aide de la fédération, vous assumez indirectement un rôle.

Selon le type d'utilisateur que vous êtes, vous pouvez vous connecter au portail AWS Management Console ou au portail AWS d'accès. Pour plus d'informations sur la connexion à AWS, consultez Comment vous connecter à votre compte Compte AWS dans le guide de Connexion à AWS l'utilisateur.

Si vous y accédez AWS par programmation, AWS fournit un kit de développement logiciel (SDK) et une interface de ligne de commande (CLI) pour signer cryptographiquement vos demandes à l'aide de vos informations d'identification. Si vous n'utilisez pas d' AWS outils, vous devez signer vous-même les demandes. Pour plus d'informations sur l'utilisation de la méthode recommandée pour signer des demandes vous-même, consultez AWS Signature Version 4 pour les demandes d'API dans le Guide de l'utilisateur IAM.

Quelle que soit la méthode d'authentification que vous utilisez, vous devrez peut-être fournir des informations de sécurité supplémentaires. Par exemple, il vous AWS recommande d'utiliser l'authentification multifactorielle (MFA) pour renforcer la sécurité de votre compte. Pour plus d'informations, consultez Authentification multifactorielle dans le Guide de l'utilisateur AWS IAM Identity Center et Authentification multifactorielle AWS dans IAM dans le Guide de l'utilisateur IAM.

Compte AWS utilisateur root

Lorsque vous créez un Compte AWS, vous commencez par une identité de connexion unique qui donne un accès complet à toutes Services AWS les ressources du compte. Cette identité est appelée utilisateur Compte AWS root et est accessible en vous connectant avec l'adresse e-mail et le mot de passe que vous avez utilisés pour créer le compte. Il est vivement recommandé de ne pas utiliser l'utilisateur racine pour vos tâches quotidiennes. Protégez vos informations d'identification d'utilisateur racine et utilisez-les pour effectuer les tâches que seul l'utilisateur racine peut effectuer. Pour obtenir la liste complète des tâches qui vous imposent de vous connecter en tant qu'utilisateur racine, consultez Tâches nécessitant les informations d'identification de l'utilisateur racine dans le Guide de l'utilisateur IAM.

Utilisateurs et groupes IAM

Un utilisateur IAM est une identité au sein de vous Compte AWS qui possède des autorisations spécifiques pour une seule personne ou application. Dans la mesure du possible, nous vous recommandons de vous appuyer sur des informations d'identification temporaires plutôt que de créer des utilisateurs IAM ayant des informations d'identification à long terme telles que des mots de passe et des clés d'accès. Toutefois, si certains cas d'utilisation spécifiques nécessitent des informations d'identification à long terme avec les utilisateurs IAM, nous vous recommandons d'effectuer une rotation des clés d'accès. Pour plus d'informations, consultez Rotation régulière des clés d'accès pour les cas d'utilisation nécessitant des informations d'identification dans le Guide de l'utilisateur IAM.

Un groupe IAM est une identité qui concerne un ensemble d'utilisateurs IAM. Vous ne pouvez pas vous connecter en tant que groupe. Vous pouvez utiliser les groupes pour spécifier des autorisations pour plusieurs utilisateurs à la fois. Les groupes permettent de gérer plus facilement les autorisations pour de grands ensembles d'utilisateurs. Par exemple, vous pouvez nommer un groupe IAMAdminset lui donner les autorisations nécessaires pour administrer les ressources IAM.

Les utilisateurs sont différents des rôles. Un utilisateur est associé de manière unique à une personne ou une application, alors qu'un rôle est conçu pour être endossé par tout utilisateur qui en a besoin. Les utilisateurs disposent d'informations d'identification permanentes, mais les rôles fournissent des informations d'identification temporaires. Pour plus d'informations, consultez Cas d'utilisation pour les utilisateurs IAM dans le Guide de l'utilisateur IAM.

Rôles IAM

Un rôle IAM est une identité au sein de vous Compte AWS dotée d'autorisations spécifiques. Le concept ressemble à celui d'utilisateur IAM, mais le rôle IAM n'est pas associé à une personne en particulier. Pour assumer temporairement un rôle IAM dans le AWS Management Console, vous pouvez passer d'un rôle d'utilisateur à un rôle IAM (console). Vous pouvez assumer un rôle en appelant une opération d' AWS API AWS CLI ou en utilisant une URL personnalisée. Pour plus d'informations sur les méthodes d'utilisation des rôles, consultez Méthodes pour endosser un rôle dans le Guide de l'utilisateur IAM.

Les rôles IAM avec des informations d'identification temporaires sont utiles dans les cas suivants :

- Accès utilisateur fédéré : pour attribuer des autorisations à une identité fédérée, vous créez un rôle et définissez des autorisations pour le rôle. Quand une identité externe s'authentifie, l'identité est associée au rôle et reçoit les autorisations qui sont définies par celui-ci. Pour obtenir des informations sur les rôles pour la fédération, consultez Création d'un rôle pour un fournisseur d'identité tiers (fédération) dans le Guide de l'utilisateur IAM. Si vous utilisez IAM Identity Center, vous configurez un jeu d'autorisations. IAM Identity Center met en corrélation le jeu d'autorisations avec un rôle dans IAM afin de contrôler à quoi vos identités peuvent accéder après leur authentification. Pour plus d'informations sur les jeux d'autorisations, consultez Jeux d'autorisations dans le Guide de l'utilisateur AWS IAM Identity Center .
- Autorisations d'utilisateur IAM temporaires : un rôle ou un utilisateur IAM peut endosser un rôle IAM pour profiter temporairement d'autorisations différentes pour une tâche spécifique.
- Accès intercompte : vous pouvez utiliser un rôle IAM pour permettre à un utilisateur (principal de confiance) d'un compte différent d'accéder aux ressources de votre compte. Les rôles constituent le principal moyen d'accorder l'accès intercompte. Toutefois, dans certains Services AWS cas, vous pouvez associer une politique directement à une ressource (au lieu d'utiliser un rôle comme proxy). Pour en savoir plus sur la différence entre les rôles et les politiques basées sur les ressources pour l'accès intercompte, consultez Accès intercompte aux ressources dans IAM dans le Guide de l'utilisateur IAM.
- Accès multiservices — Certains Services AWS utilisent des fonctionnalités dans d'autres Services AWS. Par exemple, lorsque vous effectuez un appel dans un service, il est courant que ce service exécute des applications dans Amazon EC2 ou stocke des objets dans Amazon S3. Un service peut le faire en utilisant les autorisations d'appel du principal, un rôle de service ou un rôle lié au service.
  - Sessions d'accès direct (FAS) : lorsque vous utilisez un utilisateur ou un rôle IAM pour effectuer des actions AWS, vous êtes considéré comme un mandant. Lorsque vous utilisez certains

services, vous pouvez effectuer une action qui initie une autre action dans un autre service. FAS utilise les autorisations du principal appelant et Service AWS, associées Service AWS à la demande, pour adresser des demandes aux services en aval. Les demandes FAS ne sont effectuées que lorsqu'un service reçoit une demande qui nécessite des interactions avec d'autres personnes Services AWS ou des ressources pour être traitée. Dans ce cas, vous devez disposer d'autorisations nécessaires pour effectuer les deux actions. Pour plus de détails sur une politique lors de la formulation de demandes FAS, consultez Transmission des sessions d'accès.

- Rôle de service : il s'agit d'un rôle IAM attribué à un service afin de réaliser des actions en votre nom. Un administrateur IAM peut créer, modifier et supprimer un rôle de service à partir d'IAM. Pour plus d'informations, consultez Création d'un rôle pour la délégation d'autorisations à un Service AWS dans le Guide de l'utilisateur IAM.
- Rôle lié à un service — Un rôle lié à un service est un type de rôle de service lié à un. Service AWS Le service peut endosser le rôle afin d'effectuer une action en votre nom. Les rôles liés au service apparaissent dans votre Compte AWS fichier et appartiennent au service. Un administrateur IAM peut consulter, mais ne peut pas modifier, les autorisations concernant les rôles liés à un service.

- Applications exécutées sur Amazon EC2 : vous pouvez utiliser un rôle IAM pour gérer les informations d'identification temporaires pour les applications qui s'exécutent sur une EC2 instance et qui envoient des demandes AWS CLI d' AWS API. Cela est préférable au stockage des clés d'accès dans l' EC2 instance. Pour attribuer un AWS rôle à une EC2 instance et le rendre disponible pour toutes ses applications, vous devez créer un profil d'instance attaché à l'instance. Un profil d'instance contient le rôle et permet aux programmes exécutés sur l' EC2 instance d'obtenir des informations d'identification temporaires. Pour plus d'informations, consultez Utiliser un rôle IAM pour accorder des autorisations aux applications exécutées sur des EC2 instances Amazon dans le guide de l'utilisateur IAM.

Gestion des accès à l'aide de politiques

Vous contrôlez l'accès en AWS créant des politiques et en les associant à AWS des identités ou à des ressources. Une politique est un objet AWS qui, lorsqu'il est associé à une identité ou à une ressource, définit leurs autorisations. AWS évalue ces politiques lorsqu'un principal (utilisateur, utilisateur root ou session de rôle) fait une demande. Les autorisations dans les politiques déterminent si la demande est autorisée ou refusée. La plupart des politiques sont stockées AWS sous forme de documents JSON. Pour plus d'informations sur la structure et le contenu des documents de politique JSON, consultez Vue d'ensemble des politiques JSON dans le Guide de l'utilisateur IAM.

Les administrateurs peuvent utiliser les politiques AWS JSON pour spécifier qui a accès à quoi. C'est-à-dire, quel principal peut effectuer des actions sur quelles ressources et dans quelles conditions.

Par défaut, les utilisateurs et les rôles ne disposent d'aucune autorisation. Pour octroyer aux utilisateurs des autorisations d'effectuer des actions sur les ressources dont ils ont besoin, un administrateur IAM peut créer des politiques IAM. L'administrateur peut ensuite ajouter les politiques IAM aux rôles et les utilisateurs peuvent assumer les rôles.

Les politiques IAM définissent les autorisations d'une action, quelle que soit la méthode que vous utilisez pour exécuter l'opération. Par exemple, supposons que vous disposiez d'une politique qui autorise l'action `iam:GetRole`. Un utilisateur appliquant cette politique peut obtenir des informations sur le rôle à partir de AWS Management Console AWS CLI, de ou de l' AWS API.

Politiques basées sur l'identité

Les politiques basées sur l'identité sont des documents de politique d'autorisations JSON que vous pouvez attacher à une identité telle qu'un utilisateur, un groupe d'utilisateurs ou un rôle IAM. Ces politiques contrôlent quel type d'actions des utilisateurs et des rôles peuvent exécuter, sur quelles ressources et dans quelles conditions. Pour découvrir comment créer une politique basée sur l'identité, consultez Définition d'autorisations IAM personnalisées avec des politiques gérées par le client dans le Guide de l'utilisateur IAM.

Les politiques basées sur l'identité peuvent être classées comme des politiques en ligne ou des politiques gérées. Les politiques en ligne sont intégrées directement à un utilisateur, groupe ou rôle. Les politiques gérées sont des politiques autonomes que vous pouvez associer à plusieurs utilisateurs, groupes et rôles au sein de votre Compte AWS. Les politiques gérées incluent les politiques AWS gérées et les politiques gérées par le client. Pour découvrir comment choisir entre une politique gérée et une politique en ligne, consultez Choix entre les politiques gérées et les politiques en ligne dans le Guide de l'utilisateur IAM.

Politiques basées sur les ressources

Les politiques basées sur les ressources sont des documents de politique JSON que vous attachez à une ressource. Par exemple, les politiques de confiance de rôle IAM et les politiques de compartiment Amazon S3 sont des politiques basées sur les ressources. Dans les services qui sont compatibles avec les politiques basées sur les ressources, les administrateurs de service peuvent les utiliser pour contrôler l'accès à une ressource spécifique. Pour la ressource dans laquelle se trouve la politique, cette dernière définit quel type d'actions un principal spécifié peut effectuer sur cette

ressource et dans quelles conditions. Vous devez spécifier un principal dans une politique basée sur les ressources. Les principaux peuvent inclure des comptes, des utilisateurs, des rôles, des utilisateurs fédérés ou. Services AWS

Les politiques basées sur les ressources sont des politiques en ligne situées dans ce service. Vous ne pouvez pas utiliser les politiques AWS gérées par IAM dans une stratégie basée sur les ressources.

Listes de contrôle d'accès (ACLs)

Les listes de contrôle d'accès (ACLs) contrôlent les principaux (membres du compte, utilisateurs ou rôles) autorisés à accéder à une ressource. ACLs sont similaires aux politiques basées sur les ressources, bien qu'elles n'utilisent pas le format de document de politique JSON.

Amazon S3 et AWS WAF Amazon VPC sont des exemples de services compatibles. ACLs Pour en savoir plus ACLs, consultez la présentation de la liste de contrôle d'accès (ACL) dans le guide du développeur Amazon Simple Storage Service.

Autres types de politique

AWS prend en charge d'autres types de politiques moins courants. Ces types de politiques peuvent définir le nombre maximum d’autorisations qui vous sont accordées par des types de politiques plus courants.

- Limite d’autorisations : une limite d’autorisations est une fonctionnalité avancée dans laquelle vous définissez le nombre maximal d’autorisations qu’une politique basée sur l’identité peut accorder à une entité IAM (utilisateur ou rôle IAM). Vous pouvez définir une limite d’autorisations pour une entité. Les autorisations en résultant représentent la combinaison des politiques basées sur l’identité d’une entité et de ses limites d’autorisation. Les politiques basées sur les ressources qui spécifient l’utilisateur ou le rôle dans le champ `Principal` ne sont pas limitées par les limites d’autorisations. Un refus explicite dans l’une de ces politiques annule l’autorisation. Pour plus d’informations sur les limites d’autorisations, consultez Limites d’autorisations pour des entités IAM dans le Guide de l’utilisateur IAM.
- Politiques de contrôle des services (SCPs) : SCPs politiques JSON qui spécifient les autorisations maximales pour une organisation ou une unité organisationnelle (UO) dans AWS Organizations. AWS Organizations est un service permettant de regrouper et de gérer de manière centralisée Comptes AWS les multiples propriétés de votre entreprise. Si vous activez toutes les fonctionnalités d'une organisation, vous pouvez appliquer des politiques de contrôle des services (SCPs) à l'un ou à l'ensemble de vos comptes. Le SCP limite les autorisations pour les entités

figurant dans les comptes des membres, y compris chacune Utilisateur racine d'un compte AWS d'entre elles. Pour plus d'informations sur les Organizations SCPs, voir Politiques de contrôle des services dans le Guide de AWS Organizations l'utilisateur.

- Politiques de contrôle des ressources (RCPs) : RCPs politiques JSON que vous pouvez utiliser pour définir le maximum d'autorisations disponibles pour les ressources de vos comptes sans mettre à jour les politiques IAM associées à chaque ressource que vous possédez. Le RCP limite les autorisations pour les ressources des comptes membres et peut avoir un impact sur les autorisations effectives pour les identités, y compris Utilisateur racine d'un compte AWS, qu'elles appartiennent ou non à votre organisation. Pour plus d'informations sur les Organizations RCPs, y compris une liste de ces Services AWS supports RCPs, consultez la section Resource control policies (RCPs) dans le guide de AWS Organizations l'utilisateur.
- Politiques de séance : les politiques de séance sont des politiques avancées que vous utilisez en tant que paramètre lorsque vous créez par programmation une séance temporaire pour un rôle ou un utilisateur fédéré. Les autorisations de séance en résultant sont une combinaison des politiques basées sur l'identité de l'utilisateur ou du rôle et des politiques de séance. Les autorisations peuvent également provenir d'une politique basée sur les ressources. Un refus explicite dans l'une de ces politiques annule l'autorisation. Pour plus d'informations, consultez Politiques de session dans le Guide de l'utilisateur IAM.

Plusieurs types de politique

Lorsque plusieurs types de politiques s'appliquent à la requête, les autorisations en résultant sont plus compliquées à comprendre. Pour savoir comment AWS déterminer s'il faut autoriser une demande lorsque plusieurs types de politiques sont impliqués, consultez la section Logique d'évaluation des politiques dans le guide de l'utilisateur IAM.

Comment Amazon Kendra travaille avec IAM

Avant d'utiliser IAM pour gérer l'accès à Amazon Kendra, vous devez connaître les fonctionnalités IAM disponibles avec Amazon Kendra. Pour obtenir une vue d'ensemble de la manière dont Amazon Kendra et les autres AWS services fonctionnent avec IAM, consultez la section AWS Services That Work with IAM dans le guide de l'utilisateur IAM.

Rubriques

- Politiques basées sur l'identité d'Amazon Kendra
- Politiques basées sur les ressources Amazon Kendra

- Listes de contrôle d'accès (ACLs)
- Autorisation basée sur les tags Amazon Kendra
- Rôles IAM d'Amazon Kendra

Politiques basées sur l'identité d'Amazon Kendra

Avec les politiques IAM basées sur l'identité, vous pouvez spécifier des actions et ressources autorisées ou refusées, ainsi que les conditions dans lesquelles les actions sont autorisées ou refusées. Amazon Kendra prend en charge des actions, des ressources et des clés de condition spécifiques. Pour en savoir plus sur tous les éléments que vous utilisez dans une politique JSON, consultez Références des éléments de politique JSON IAM dans le Guide de l'utilisateur IAM.

Actions

Les administrateurs peuvent utiliser les politiques AWS JSON pour spécifier qui a accès à quoi. C'est-à-dire, quel principal peut effectuer des actions sur quelles ressources et dans quelles conditions.

L'élément `Action` d'une politique JSON décrit les actions que vous pouvez utiliser pour autoriser ou refuser l'accès à une politique. Les actions de stratégie portent généralement le même nom que l'opération AWS d'API associée. Il existe quelques exceptions, telles que les actions avec autorisations uniquement qui n'ont pas d'opération API correspondante. Certaines opérations nécessitent également plusieurs actions dans une politique. Ces actions supplémentaires sont nommées actions dépendantes.

Intégration d'actions dans une politique afin d'accorder l'autorisation d'exécuter les opérations associées.

Les actions politiques dans Amazon Kendra utilisent le préfixe suivant avant l'action :. `kendra:` Par exemple, pour autoriser une personne à répertorier les index Amazon Kendra avec l'opération d'ListIndicesAPI, vous devez inclure cette `kendra:ListIndices` action dans sa politique. Les déclarations de politique doivent inclure un élément `Action` ou `NotAction`. Amazon Kendra définit son propre ensemble d'actions décrivant les tâches que vous pouvez effectuer avec ce service.

Pour spécifier plusieurs actions dans une seule déclaration, séparez-les par des virgules comme suit :

```
"Action": [
      "kendra:action1",
```

```
      "kendra:action2"
```

Vous pouvez aussi spécifier plusieurs actions à l'aide de caractères génériques (*). Par exemple, pour spécifier toutes les actions qui commencent par le mot `Describe`, incluez l'action suivante :

```
"Action": "kendra:Describe*"
```

Pour consulter la liste des actions Amazon Kendra, consultez la section Actions définies par Amazon Kendra dans le guide de l'utilisateur IAM.

Ressources

Les administrateurs peuvent utiliser les politiques AWS JSON pour spécifier qui a accès à quoi. C'est-à-dire, quel principal peut effectuer des actions sur quelles ressources et dans quelles conditions.

L'élément de politique JSON `Resource` indique le ou les objets auxquels l'action s'applique. Les instructions doivent inclure un élément `Resource` ou `NotResource`. Il est recommandé de définir une ressource à l'aide de son Amazon Resource Name (ARN). Vous pouvez le faire pour des actions qui prennent en charge un type de ressource spécifique, connu sous la dénomination autorisations de niveau ressource.

Pour les actions qui ne sont pas compatibles avec les autorisations de niveau ressource, telles que les opérations de liste, utilisez un caractère générique (*) afin d'indiquer que l'instruction s'applique à toutes les ressources.

```
"Resource": "*"
```

La ressource d'index Amazon Kendra possède l'ARN suivant :

```
arn:${Partition}:kendra:${Region}:${Account}:index/${IndexId}
```

Pour plus d'informations sur le format de ARNs, consultez Amazon Resource Names (ARNs) et AWS Service Namespaces.

Par exemple, pour spécifier un index dans votre instruction, utilisez le GUID de l'index dans l'ARN suivant :

```
"Resource": "arn:aws:kendra:${Region}:${Account}:index/${GUID}"
```

Pour spécifier tous les index appartenant à un compte spécifique, utilisez le caractère générique (*) :

```
"Resource": "arn:aws:${Region}:${Account}:index/*"
```

Certaines actions Amazon Kendra, telles que celles relatives à la création de ressources, ne peuvent pas être effectuées sur une ressource spécifique. Dans ces cas-là, vous devez utiliser le caractère générique (*).

```
"Resource": "*"
```

Pour consulter la liste des types de ressources Amazon Kendra et leurs caractéristiques ARNs, consultez la section Ressources définies par Amazon Kendra dans le guide de l'utilisateur IAM. Pour savoir avec quelles actions vous pouvez spécifier l'ARN de chaque ressource, consultez Actions définies par Amazon Kendra.

Clés de condition

Les administrateurs peuvent utiliser les politiques AWS JSON pour spécifier qui a accès à quoi. C'est-à-dire, quel principal peut effectuer des actions sur quelles ressources et dans quelles conditions.

L'élément `Condition` (ou le bloc `Condition`) vous permet de spécifier des conditions lorsqu'une instruction est appliquée. L'élément `Condition` est facultatif. Vous pouvez créer des expressions conditionnelles qui utilisent des opérateurs de condition, tels que les signes égal ou inférieur à, pour faire correspondre la condition de la politique aux valeurs de la demande.

Si vous spécifiez plusieurs éléments `Condition` dans une instruction, ou plusieurs clés dans un seul élément `Condition`, AWS les évalue à l'aide d'une opération `AND` logique. Si vous spécifiez plusieurs valeurs pour une seule clé de condition, AWS évalue la condition à l'aide d'une `OR` opération logique. Toutes les conditions doivent être remplies avant que les autorisations associées à l'instruction ne soient accordées.

Vous pouvez aussi utiliser des variables d'espace réservé quand vous spécifiez des conditions. Par exemple, vous pouvez accorder à un utilisateur IAM l'autorisation d'accéder à une ressource uniquement si elle est balisée avec son nom d'utilisateur IAM. Pour plus d'informations, consultez Éléments d'une politique IAM : variables et identifications dans le Guide de l'utilisateur IAM.

AWS prend en charge les clés de condition globales et les clés de condition spécifiques au service. Pour voir toutes les clés de condition AWS globales, voir les clés de contexte de condition AWS globales dans le guide de l'utilisateur IAM.

Amazon Kendra ne fournit aucune clé de condition spécifique à un service, mais prend en charge l'utilisation de certaines clés de condition globales. Pour voir toutes les clés de condition AWS globales, consultez la section Clés contextuelles de condition AWS globale dans le guide de l'utilisateur IAM.

Exemples

Pour consulter des exemples de politiques basées sur l'identité d'Amazon Kendra, consultez. Exemples de politiques basées sur l'identité d'Amazon Kendra

Politiques basées sur les ressources Amazon Kendra

Amazon Kendra ne prend pas en charge les politiques basées sur les ressources.

Listes de contrôle d'accès (ACLs)

Amazon Kendra ne prend pas en charge les listes de contrôle d'accès (ACLs) pour l'accès aux AWS services et aux ressources.

Autorisation basée sur les tags Amazon Kendra

Vous pouvez associer des balises à certains types de ressources Amazon Kendra pour autoriser l'accès à ces ressources. Pour contrôler l'accès en fonction des balises, fournissez les informations relatives aux balises dans l'élément de condition d'une politique à l'aide des clés`aws:RequestTag/key-name`, ou de `aws:TagKeys` condition.

Le tableau suivant répertorie les actions, les types de ressources correspondants et les clés de condition pour le contrôle d'accès basé sur des balises. Chaque action est autorisée en fonction des balises associées au type de ressource correspondant.

Action	Type de ressource	Clés de condition
CreateDataSource		`aws:RequestTag` , `aws:TagKeys`
CreateFaq		`aws:RequestTag` , `aws:TagKeys`
CreateIndex		`aws:RequestTag` , `aws:TagKeys`

Action	Type de ressource	Clés de condition
API_ListTagsForRessource	source de données, FAQ, index	
TagResource	source de données, FAQ, index	`aws:RequestTag` , `aws:TagKeys`
UntagResource	source de données, FAQ, index	`aws:TagKeys`

Pour plus d'informations sur le balisage des ressources Amazon Kendra, consultez. Balises Pour un exemple de politique basée sur l'identité qui limite l'accès à une ressource en fonction des balises de ressource, voir. Exemples de politique basée sur des balises Pour plus d'informations sur l'utilisation de balises pour limiter l'accès aux ressources, consultez la section Contrôle de l'accès à l'aide de balises dans le guide de l'utilisateur IAM.

Rôles IAM d'Amazon Kendra

Un rôle IAM est une entité de votre AWS compte qui dispose d'autorisations spécifiques.

Utilisation d'informations d'identification temporaires avec Amazon Kendra

Vous pouvez utiliser des informations d’identification temporaires pour vous connecter à l’aide de la fédération, endosser un rôle IAM ou encore pour endosser un rôle intercompte. Vous obtenez des informations d'identification de sécurité temporaires en appelant des opérations d' AWS STS API telles que AssumeRoleou GetFederationToken.

Amazon Kendra prend en charge l'utilisation d'informations d'identification temporaires.

Rôles de service

Cette fonction permet à un service d’endosser une fonction du service en votre nom. Ce rôle autorise le service à accéder à des ressources d’autres services pour effectuer une action en votre nom. Les rôles de service s’affichent dans votre compte IAM et sont la propriété du compte. Cela signifie qu’un administrateur IAM peut modifier les autorisations associées à ce rôle. Toutefois, une telle action peut perturber le bon fonctionnement du service.

Amazon Kendra prend en charge les rôles de service.

Choisir un rôle IAM dans Amazon Kendra

Lorsque vous créez un index, appelez l'`BatchPutDocument`opération, créez une source de données ou créez une FAQ, vous devez fournir un rôle d'accès Amazon Resource Name (ARN) qu'Amazon Kendra utilise pour accéder aux ressources requises en votre nom. Si vous avez déjà créé un rôle, la console Amazon Kendra vous propose une liste de rôles parmi lesquels choisir. Il est important de choisir un rôle qui permet d'accéder aux ressources dont vous avez besoin. Pour de plus amples informations, veuillez consulter IAM rôles d'accès pour Amazon Kendra.

Exemples de politiques basées sur l'identité d'Amazon Kendra

Par défaut, les utilisateurs et les rôles ne sont pas autorisés à créer ou à modifier les ressources Amazon Kendra. Ils ne peuvent pas non plus effectuer de tâches à l'aide de l' AWS API AWS Management Console AWS CLI, ou. Un administrateur IAM doit créer des politiques IAM autorisant les utilisateurs et les rôles à exécuter des opérations d’API spécifiques sur les ressources spécifiées dont ils ont besoin. Il doit ensuite attacher ces stratégies aux utilisateurs ou aux groupes ayant besoin de ces autorisations.

Pour savoir comment créer une politique IAM basée sur l’identité à l’aide de ces exemples de documents de politique JSON, consultez Création de politiques dans l’onglet JSON dans le Guide de l’utilisateur IAM.

Rubriques

- Bonnes pratiques en matière de politiques
- AWS Politiques gérées (prédéfinies) pour Amazon Kendra
- Autorisation accordée aux utilisateurs pour afficher leurs propres autorisations
- Accès à un index Amazon Kendra
- Exemples de politique basée sur des balises

Bonnes pratiques en matière de politiques

Les politiques basées sur l'identité déterminent si quelqu'un peut créer, accéder ou supprimer les ressources Amazon Kendra de votre compte. Ces actions peuvent entraîner des frais pour votre Compte AWS. Lorsque vous créez ou modifiez des politiques basées sur l’identité, suivez ces instructions et recommandations :

- Commencez AWS par les politiques gérées et passez aux autorisations du moindre privilège : pour commencer à accorder des autorisations à vos utilisateurs et à vos charges de travail, utilisez

les politiques AWS gérées qui accordent des autorisations pour de nombreux cas d'utilisation courants. Ils sont disponibles dans votre Compte AWS. Nous vous recommandons de réduire davantage les autorisations en définissant des politiques gérées par les AWS clients spécifiques à vos cas d'utilisation. Pour plus d'informations, consultez politiques gérées par AWS ou politiques gérées par AWS pour les activités professionnelles dans le Guide de l'utilisateur IAM.

- Accordez les autorisations de moindre privilège : lorsque vous définissez des autorisations avec des politiques IAM, accordez uniquement les autorisations nécessaires à l'exécution d'une seule tâche. Pour ce faire, vous définissez les actions qui peuvent être entreprises sur des ressources spécifiques dans des conditions spécifiques, également appelées autorisations de moindre privilège. Pour plus d'informations sur l'utilisation d'IAM pour appliquer des autorisations, consultez politiques et autorisations dans IAM dans le Guide de l'utilisateur IAM.
- Utilisez des conditions dans les politiques IAM pour restreindre davantage l'accès : vous pouvez ajouter une condition à vos politiques afin de limiter l'accès aux actions et aux ressources. Par exemple, vous pouvez écrire une condition de politique pour spécifier que toutes les demandes doivent être envoyées via SSL. Vous pouvez également utiliser des conditions pour accorder l'accès aux actions de service si elles sont utilisées par le biais d'un service spécifique Service AWS, tel que AWS CloudFormation. Pour plus d'informations, consultez Conditions pour éléments de politique JSON IAM dans le Guide de l'utilisateur IAM.
- Utilisez l'Analyseur d'accès IAM pour valider vos politiques IAM afin de garantir des autorisations sécurisées et fonctionnelles : l'Analyseur d'accès IAM valide les politiques nouvelles et existantes de manière à ce que les politiques IAM respectent le langage de politique IAM (JSON) et les bonnes pratiques IAM. IAM Access Analyzer fournit plus de 100 vérifications de politiques et des recommandations exploitables pour vous aider à créer des politiques sécurisées et fonctionnelles. Pour plus d'informations, consultez Validation de politiques avec IAM Access Analyzer dans le Guide de l'utilisateur IAM.
- Exiger l'authentification multifactorielle (MFA) : si vous avez un scénario qui nécessite des utilisateurs IAM ou un utilisateur root, activez l'authentification MFA pour une sécurité accrue. Compte AWS Pour exiger la MFA lorsque des opérations d'API sont appelées, ajoutez des conditions MFA à vos politiques. Pour plus d'informations, consultez Sécurisation de l'accès aux API avec MFA dans le Guide de l'utilisateur IAM.

Pour plus d'informations sur les bonnes pratiques dans IAM, consultez Bonnes pratiques de sécurité dans IAM dans le Guide de l'utilisateur IAM.

AWS Politiques gérées (prédéfinies) pour Amazon Kendra

AWS répond à de nombreux cas d'utilisation courants en fournissant des politiques IAM autonomes créées et administrées par. AWS Ces politiques sont appelées politiques AWS gérées. AWS les politiques gérées vous permettent d'attribuer plus facilement des autorisations aux utilisateurs, aux groupes et aux rôles que si vous deviez les rédiger vous-même. Pour plus d'informations, consultez Ajout d'autorisations à un utilisateur dans le Guide de l'utilisateur IAM.

Les politiques AWS gérées suivantes, que vous pouvez associer aux groupes et aux rôles de votre compte, sont spécifiques à Amazon Kendra :

- AmazonKendraReadOnly— Accorde un accès en lecture seule aux ressources Amazon Kendra.
- AmazonKendraFullAccess— Accorde un accès complet pour créer, lire, mettre à jour, supprimer, étiqueter et exécuter toutes les ressources Amazon Kendra.

Pour la console, votre rôle doit également comporter des `s3:ListBucket` autorisations `iam:CreateRole` `iam:CreatePolicyiam:AttachRolePolicy`,, et.

> **Note**
>
> Vous pouvez vérifier ces autorisations en vous connectant à la console IAM et en recherchant des politiques spécifiques.

Vous pouvez également créer vos propres politiques personnalisées pour autoriser les actions d'API Amazon Kendra. Vous pouvez attacher ces politiques personnalisées aux rôles ou groupes IAM qui nécessitent ces autorisations. Pour des exemples de politiques IAM pour Amazon Kendra, consultez. Exemples de politiques basées sur l'identité d'Amazon Kendra

Autorisation accordée aux utilisateurs pour afficher leurs propres autorisations

Cet exemple montre comment créer une politique qui permet aux utilisateurs IAM d'afficher les politiques en ligne et gérées attachées à leur identité d'utilisateur. Cette politique inclut les autorisations permettant d'effectuer cette action sur la console ou par programmation à l'aide de l'API AWS CLI or AWS .

```
{
    "Version": "2012-10-17",
    "Statement": [
```

```
        {
            "Sid": "ViewOwnUserInfo",
            "Effect": "Allow",
            "Action": [
                "iam:GetUserPolicy",
                "iam:ListGroupsForUser",
                "iam:ListAttachedUserPolicies",
                "iam:ListUserPolicies",
                "iam:GetUser"
            ],
            "Resource": ["arn:aws:iam::*:user/${aws:username}"]
        },
        {
            "Sid": "NavigateInConsole",
            "Effect": "Allow",
            "Action": [
                "iam:GetGroupPolicy",
                "iam:GetPolicyVersion",
                "iam:GetPolicy",
                "iam:ListAttachedGroupPolicies",
                "iam:ListGroupPolicies",
                "iam:ListPolicyVersions",
                "iam:ListPolicies",
                "iam:ListUsers"
            ],
            "Resource": "*"
        }
    ]
}
```

Accès à un index Amazon Kendra

Dans cet exemple, vous souhaitez autoriser un utilisateur de votre AWS compte à interroger un index.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "QueryIndex",
```

```
            "Effect": "Allow",
            "Action": [
                ":Query"
            ],
            "Resource": "arn:aws::${Region}:${Account}:index/${Index ID}"
        }
    ]
}
```

Exemples de politique basée sur des balises

Les politiques basées sur des balises sont des documents de politique JSON qui spécifient les actions qu'un principal peut effectuer sur des ressources balisées.

Exemple : utiliser un tag pour accéder à une ressource

Cet exemple de politique accorde à un utilisateur ou à un rôle de votre AWS compte l'autorisation d'utiliser l'`Query`opération avec n'importe quelle ressource associée à la clé **`department`** et à la valeur**`finance`**.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "kendra:Query"
            ],
            "Resource": "*",
            "Condition": {
                "StringEquals": {
                    "aws:ResourceTag/department": "finance"
                }
            }
        }
   ]
}
```

Exemple : utiliser un tag pour activer les opérations Amazon Kendra

Cet exemple de politique accorde à un utilisateur ou à un rôle de votre AWS compte l'autorisation d'utiliser n'importe quelle opération Amazon Kendra, à l'exception des `TagResource` opérations avec une ressource étiquetée avec la clé **`department`** et la valeur. **`finance`**

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": "kendra:*",
            "Resource": "*"
        },
        {
            "Effect": "Deny",
            "Action": [
                "kendra:TagResource"
            ],
            "Resource": "*",
            "Condition": {
                "StringEquals": {
                    "aws:ResourceTag/department": "finance"
                }
            }
        }
    ]
}
```

Exemple : utilisation d'une balise pour restreindre l'accès à une opération

Cet exemple de politique restreint l'accès d'un utilisateur ou d'un rôle de votre AWS compte pour utiliser l'`CreateIndex`opération, sauf si l'utilisateur fournit le **`department`** tag et que celui-ci possède les valeurs autorisées **`finance`** et**`IT`**.

JSON

```
{
```

```
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": "kendra:CreateIndex",
            "Resource": "*"
        },
        {
            "Effect": "Deny",
            "Action": "kendra:CreateIndex",
            "Resource": "*",
            "Condition": {
                "Null": {
                    "aws:RequestTag/department": "true"
                }
            }
        },
        {
            "Effect": "Deny",
            "Action": "kendra:CreateIndex",
            "Resource": "*",
            "Condition": {
                "ForAnyValue:StringNotEquals": {
                    "aws:RequestTag/department": [
                        "finance",
                        "IT"
                    ]
                }
            }
        }
    ]
}
```

AWS politiques gérées pour Amazon Kendra

Pour ajouter des autorisations aux utilisateurs, aux groupes et aux rôles, il est plus facile d'utiliser des politiques AWS gérées que de les rédiger vous-même. Il faut du temps et de l’expertise pour créer des politiques gérées par le client IAM qui ne fournissent à votre équipe que les autorisations dont elle a besoin. Pour démarrer rapidement, vous pouvez utiliser nos politiques AWS gérées. Ces politiques couvrent les cas d'utilisation courants et sont disponibles dans votre AWS compte. Pour

plus d'informations sur les politiques AWS gérées, voir les politiques AWS gérées dans le guide de l'utilisateur IAM.

AWS les services maintiennent et mettent à jour les politiques AWS gérées. Vous ne pouvez pas modifier les autorisations dans les politiques AWS gérées. Les services ajoutent occasionnellement des autorisations à une politique gérée par AWS pour prendre en charge de nouvelles fonctionnalités. Ce type de mise à jour affecte toutes les identités (utilisateurs, groupes et rôles) auxquelles la politique est attachée. Les services sont très susceptibles de mettre à jour une politique gérée par AWS quand une nouvelle fonctionnalité est lancée ou quand de nouvelles opérations sont disponibles. Les services ne suppriment pas les autorisations d'une politique AWS gérée. Les mises à jour des politiques n'endommageront donc pas vos autorisations existantes.

En outre, AWS prend en charge les politiques gérées pour les fonctions professionnelles qui couvrent plusieurs services. Par exemple, la politique ReadOnlyAccess AWS gérée fournit un accès en lecture seule à tous les AWS services et ressources. Lorsqu'un service lance une nouvelle fonctionnalité, il AWS ajoute des autorisations en lecture seule pour les nouvelles opérations et ressources. Pour obtenir la liste des politiques de fonctions professionnelles et leurs descriptions, consultez la page politiques gérées par AWS pour les fonctions de tâche dans le Guide de l'utilisateur IAM.

AWS politique gérée : AmazonKendraReadOnly

Accorde un accès en lecture seule aux ressources Amazon Kendra. Cette politique inclut les autorisations suivantes.

- `kendra`— Permet aux utilisateurs d'effectuer des actions qui renvoient soit une liste d'éléments, soit des détails sur un élément. Cela inclut les opérations d'API qui commencent par `Describe` `ListQuery`,`BatchGetDocumentStatus`,`GetQuerySuggestions`, ou`GetSnapshots`.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Action": [
                "kendra:Describe*",
                "kendra:List*",
```

```
                "kendra:Query",
                "kendra:BatchGetDocumentStatus",
                "kendra:GetQuerySuggestions",
                "kendra:GetSnapshots"
            ],
            "Effect": "Allow",
            "Resource": "*"
        }
    ]
}
```

AWS politique gérée : AmazonKendraFullAccess

Accorde un accès complet à la création, à la lecture, à la mise à jour, à la suppression, au balisage et à l'exécution de toutes les ressources Amazon Kendra. Cette politique inclut les autorisations suivantes.

- `kendra`—Permet aux principaux d'accéder en lecture et en écriture à toutes les actions d'Amazon Kendra.
- `s3`—Permet aux directeurs d'obtenir l'emplacement des compartiments Amazon S3 et de répertorier les compartiments.
- `iam`—Permet aux directeurs de transmettre et de répertorier les rôles.
- `kms`—Permet aux directeurs de décrire et de répertorier les AWS KMS clés et les alias.
- `secretsmanager`—Permet aux principaux de créer, de décrire et de répertorier des secrets.
- `ec2`—Permet aux principaux de décrire les groupes de sécurité VCPs (Virtual Private Cloud) et les sous-réseaux.
- `cloudwatch`—Permet aux administrateurs de consulter les statistiques de Cloud Watch.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": "iam:PassRole",
            "Resource": "*",
```

```
            "Condition": {
                "StringEquals": {
                    "iam:PassedToService": "kendra.amazonaws.com"
                }
            }
        },
        {
            "Effect": "Allow",
            "Action": [
                "iam:ListRoles"
            ],
            "Resource": "*"
        },
        {
            "Effect": "Allow",
            "Action": [
                "ec2:DescribeSecurityGroups",
                "ec2:DescribeVpcs",
                "ec2:DescribeSubnets"
            ],
            "Resource": "*"
        },
        {
            "Effect": "Allow",
            "Action": [
                "kms:ListKeys",
                "kms:ListAliases",
                "kms:DescribeKey"
            ],
            "Resource": "*"
        },
        {
            "Effect": "Allow",
            "Action": [
                "s3:ListAllMyBuckets",
                "s3:GetBucketLocation"
            ],
            "Resource": "*"
        },
        {
            "Effect": "Allow",
            "Action": [
                "secretsmanager:ListSecrets"
            ],
```

```
            "Resource": "*"
        },
        {
            "Effect": "Allow",
            "Action": [
                "cloudwatch:GetMetricData"
            ],
            "Resource": "*"
        },
        {
            "Effect": "Allow",
            "Action": [
                "secretsmanager:CreateSecret",
                "secretsmanager:DescribeSecret"
            ],
            "Resource": "arn:aws:secretsmanager:*:*:secret:AmazonKendra-*"
        },
        {
            "Effect": "Allow",
            "Action": "kendra:*",
            "Resource": "*"
        }
    ]
}
```

Amazon Kendra met à jour ses politiques gérées AWS

Consultez les informations relatives aux mises à jour des politiques AWS gérées pour Amazon Kendra depuis que ce service a commencé à suivre ces modifications. Pour recevoir des alertes automatiques concernant les modifications apportées à cette page, abonnez-vous au flux RSS sur la page d'historique des documents Amazon Kendra.

Modification	Description	Date
AmazonKendraReadOnly—Ajoutez l'autorisation de support GetSnapshots,	Amazon Kendra a ajouté de nouveaux APIs `GetSnapshots` et. `BatchGetD`	3 janvier 2022

Modification	Description	Date
BatchGetDocumentStatus APIs	`ocumentStatus GetSnapshots` fournit des données qui montrent comment vos utilisateurs interagissent avec votre application de recherche. `BatchGetDocumentSt atus` surveille la progressi on de l'indexation de vos documents.	
AmazonKendraReadOnly— Ajouter une autorisation pour soutenir l'opération GetQueryS uggestions	Amazon Kendra a ajouté une nouvelle API `GetQueryS uggestions` qui permet d'accéder à des suggestio ns de requêtes pour les requêtes de recherche les plus populaires, afin de guider les recherches de vos utilisateurs. Lorsque les utilisateurs saisissent leur requête de recherche, la requête suggérée les aide à compléter automatiquement leur recherche.	27 mai 2021
Amazon Kendra a commencé à suivre les modifications	Amazon Kendra a commencé à suivre les modifications apportées à ses politiques AWS gérées.	27 mai 2021

Résolution des problèmes liés à l'identité et à l'accès à Amazon Kendra

Utilisez les informations suivantes pour vous aider à diagnostiquer et à résoudre les problèmes courants que vous pouvez rencontrer lorsque vous travaillez avec Amazon Kendra et IAM.

Rubriques

- Je ne suis pas autorisé à effectuer une action dans Amazon Kendra
- Je ne suis pas autorisé à effectuer iam : PassRole
- Je suis administrateur et je souhaite autoriser d'autres personnes à accéder à Amazon Kendra
- Je souhaite autoriser des personnes extérieures à mon AWS compte à accéder à mes ressources Amazon Kendra

Je ne suis pas autorisé à effectuer une action dans Amazon Kendra

S'il vous AWS Management Console indique que vous n'êtes pas autorisé à effectuer une action, vous devez contacter votre administrateur pour obtenir de l'aide. Votre administrateur est la personne qui vous a fourni vos informations de connexion.

L'exemple d'erreur suivant se produit lorsque l'`mateojackson`utilisateur essaie d'utiliser la console pour afficher les détails d'un index mais ne dispose pas des `kendra:DescribeIndex` autorisations nécessaires.

```
User: arn:aws:iam::123456789012:user/mateojackson is not authorized to perform:
 kendra:DescribeIndex on resource: index ARN
```

Dans ce cas, Mateo demande à son administrateur de mettre à jour ses politiques pour lui permettre d'accéder à la ressource `index` à l'aide de l'action `kendra:DescribeIndex`.

Je ne suis pas autorisé à effectuer iam : PassRole

Si vous recevez un message d'erreur indiquant que vous n'êtes pas autorisé à effectuer l'`iam:PassRole`action, vos politiques doivent être mises à jour pour vous permettre de transmettre un rôle à Amazon Kendra.

Certains vous Services AWS permettent de transmettre un rôle existant à ce service au lieu de créer un nouveau rôle de service ou un rôle lié à un service. Pour ce faire, un utilisateur doit disposer des autorisations nécessaires pour transmettre le rôle au service.

L'exemple d'erreur suivant se produit lorsqu'un utilisateur IAM nommé `marymajor` essaie d'utiliser la console pour effectuer une action dans Amazon Kendra. Toutefois, l'action nécessite que le service ait des autorisations accordées par un rôle de service. Mary ne dispose pas des autorisations nécessaires pour transférer le rôle au service.

```
User: arn:aws:iam::123456789012:user/marymajor is not authorized to perform:
 iam:PassRole
```

Dans ce cas, les politiques de Mary doivent être mises à jour pour lui permettre d'exécuter l'action `iam:PassRole`.

Si vous avez besoin d'aide, contactez votre AWS administrateur. Votre administrateur vous a fourni vos informations d'identification de connexion.

Je suis administrateur et je souhaite autoriser d'autres personnes à accéder à Amazon Kendra

Pour autoriser d'autres personnes à accéder à Amazon Kendra, vous devez autoriser les personnes ou les applications qui ont besoin d'y accéder. Si vous utilisez AWS IAM Identity Center pour gérer des personnes et des applications, vous attribuez des ensembles d'autorisations aux utilisateurs ou aux groupes afin de définir leur niveau d'accès. Les ensembles d'autorisations créent et attribuent automatiquement des politiques IAM aux rôles IAM associés à la personne ou à l'application. Pour plus d'informations, consultez la section Ensembles d'autorisations dans le guide de AWS IAM Identity Center l'utilisateur.

Si vous n'utilisez pas IAM Identity Center, vous devez créer des entités IAM (utilisateurs ou rôles) pour les personnes ou les applications qui ont besoin d'un accès. Vous devez ensuite joindre une politique à l'entité qui lui accorde les autorisations appropriées dans Amazon Kendra. Une fois les autorisations accordées, fournissez les informations d'identification à l'utilisateur ou au développeur de l'application. Ils utiliseront ces informations d'identification pour y accéder AWS. Pour en savoir plus sur la création d'utilisateurs, de groupes, de politiques et d'autorisations IAM, consultez la section Identités, politiques et autorisations IAM dans le guide de l'utilisateur IAM.

Je souhaite autoriser des personnes extérieures à mon AWS compte à accéder à mes ressources Amazon Kendra

Vous pouvez créer un rôle que les utilisateurs provenant d'autres comptes ou les personnes extérieures à votre organisation pourront utiliser pour accéder à vos ressources. Vous pouvez spécifier qui est autorisé à assumer le rôle. Pour les services qui prennent en charge les politiques basées sur les ressources ou les listes de contrôle d'accès (ACLs), vous pouvez utiliser ces politiques pour autoriser les utilisateurs à accéder à vos ressources.

Pour plus d'informations, consultez les éléments suivants :

- Pour savoir si Amazon Kendra prend en charge ces fonctionnalités, consultez. Comment Amazon Kendra travaille avec IAM
- Pour savoir comment fournir l'accès à vos ressources sur celles Comptes AWS que vous possédez, consultez la section Fournir l'accès à un utilisateur IAM dans un autre utilisateur Compte AWS que vous possédez dans le Guide de l'utilisateur IAM.
- Pour savoir comment fournir l'accès à vos ressources à des tiers Comptes AWS, consultez la section Fournir un accès à des ressources Comptes AWS détenues par des tiers dans le guide de l'utilisateur IAM.
- Pour savoir comment fournir un accès par le biais de la fédération d'identité, consultez Fournir un accès à des utilisateurs authentifiés en externe (fédération d'identité) dans le Guide de l'utilisateur IAM.
- Pour en savoir plus sur la différence entre l'utilisation des rôles et des politiques basées sur les ressources pour l'accès intercompte, consultez Accès intercompte aux ressources dans IAM dans le Guide de l'utilisateur IAM.

Bonnes pratiques de sécurité

Amazon Kendra propose un certain nombre de fonctionnalités de sécurité à prendre en compte lors de l'élaboration et de la mise en œuvre de vos propres politiques de sécurité. Les bonnes pratiques suivantes doivent être considérées comme des instructions générales et ne représentent pas une solution de sécurité complète. Étant donné que ces bonnes pratiques peuvent ne pas être appropriées ou suffisantes pour votre environnement, considérez-les comme des remarques utiles plutôt que comme des recommandations.

Application du principe du moindre privilège

Amazon Kendra fournit une politique d'accès précise pour les applications utilisant des rôles. IAM Nous recommandons de n'accorder aux rôles que l'ensemble minimal de privilèges requis par le travail, tels que la couverture de votre candidature et l'accès à la destination du journal. Nous recommandons également de vérifier régulièrement les autorisations des tâches et lors de toute modification de votre application.

Autorisations de contrôle d'accès basé sur les rôles (RBAC)

Les administrateurs doivent contrôler strictement les autorisations de contrôle d'accès basé sur les rôles (RBAC) pour les applications Amazon Kendra.

Journalisation et surveillance dans Amazon Kendra

La surveillance joue un rôle important dans le maintien de la fiabilité, de la disponibilité et des performances de vos applications Amazon Kendra. Pour surveiller les appels d'API Amazon Kendra, vous pouvez utiliser. AWS CloudTrail Pour suivre le statut de vos tâches, utilisez Amazon CloudWatch Logs.

- Amazon CloudWatch Alarms : à l'aide des CloudWatch alarmes, vous surveillez une seule métrique sur une période que vous spécifiez. Si la métrique dépasse une politique. CloudWatch les alarmes n'appellent aucune action lorsqu'une métrique est dans un état particulier. L'état doit avoir changé et avoir été conservé pendant un nombre de périodes spécifié. Pour de plus amples informations, veuillez consulter Surveiller Amazon Kendra avec Amazon CloudWatch.
- AWS CloudTrail Logs : CloudTrail fournit un enregistrement des actions entreprises par un utilisateur, un rôle ou un AWS service dans Amazon Kendra ou Amazon Kendra Intelligent Ranking. À l'aide des informations collectées par CloudTrail, vous pouvez déterminer la demande qui a été faite à Amazon Kendra, l'adresse IP à partir de laquelle la demande a été faite, l'auteur de la demande, la date à laquelle elle a été faite, ainsi que des informations supplémentaires. Pour plus d'informations, consultez Enregistrement des appels d'API Amazon Kendra avec des journaux AWS CloudTrail et Enregistrement des appels de l'API Amazon Kendra Intelligent Ranking avec des journaux AWS CloudTrail.

Validation de conformité pour Amazon Kendra

Des auditeurs tiers évaluent la sécurité et la conformité d'Amazon Kendra dans le cadre de plusieurs programmes de conformité d'Amazon Kendra. Amazon Kendra est conforme aux normes suivantes :

- Health Insurance Portability and Accountability Act (HIPAA)
- Contrôles du système et de l'organisation (SOC) 2
- Programme d'évaluateurs agréés en sécurité de l'information (IRAP)
- Programme fédéral de gestion des risques et des autorisations (FedRAMP) modéré dans les régions des États-Unis East/West
- Programme fédéral de gestion des risques et des autorisations (FedRAMP) à un niveau élevé dans la région GovCloud AWS (USA Ouest)

Pour une liste des AWS services concernés par des programmes de conformité spécifiques, voir AWS Services concernés par programme de conformitéAWS . Pour des informations générales, voir Programmes de AWS conformité Programmes AWS de .

Vous pouvez télécharger des rapports d'audit tiers à l'aide de AWS Artifact. Pour plus d'informations, voir Téléchargement de rapports dans AWS Artifact .

Lorsque vous utilisez Amazon Kendra, votre responsabilité en matière de conformité dépend de la sensibilité de vos données, des objectifs de conformité de votre entreprise et des lois et réglementations applicables. AWS fournit les ressources suivantes pour faciliter la mise en conformité :

- Guides de démarrage rapide sur la sécurité et la conformité Guides sur la sécurité et la conformité : ces guides de déploiement abordent les considérations architecturales et indiquent les étapes à suivre pour déployer des environnements de base axés sur la sécurité et la conformité sur. AWS
- Livre blanc sur l'architecture pour la sécurité et la conformité HIPAA : ce livre blanc décrit comment les entreprises peuvent créer des applications conformes à la loi HIPAA. AWS
- AWS Ressources de conformité Ressources de : cette collection de classeurs et de guides peut s'appliquer à votre secteur d'activité et à votre région.
- Évaluation des ressources à l'aide des règles du guide du AWS Config développeur : le AWS Config service évalue dans quelle mesure les configurations de vos ressources sont conformes aux pratiques internes, aux directives du secteur et aux réglementations.
- AWS Security Hub—Ce AWS service fournit une vue complète de l'état de votre sécurité interne, AWS ce qui vous permet de vérifier votre conformité aux normes et aux meilleures pratiques du secteur de la sécurité.

La résilience chez Amazon Kendra

L'infrastructure AWS mondiale est construite autour des AWS régions et des zones de disponibilité. AWS Les régions fournissent plusieurs zones de disponibilité physiquement séparées et isolées, connectées par un réseau à faible latence, à haut débit et hautement redondant. Avec les zones de disponibilité, vous pouvez concevoir et exploiter des applications et des bases de données qui basculent automatiquement d’une zone à l’autre sans interruption. Les zones de disponibilité sont davantage disponibles, tolérantes aux pannes et ont une plus grande capacité de mise à l’échelle que les infrastructures traditionnelles à un ou plusieurs centres de données.

Pour plus d'informations sur AWS les régions et les zones de disponibilité, consultez la section Infrastructure AWS mondiale.

Grâce à AWS son infrastructure mondiale, Amazon Kendra Enterprise Edition est tolérant aux pannes, évolutif et hautement disponible. Le retour aux versions précédentes d'un index n'est actuellement pas pris en charge, mais vous pouvez actualiser ou recréer des parties de votre index en supprimant des sources de données existantes et en y ajoutant à nouveau des sources de données existantes.

Sécurité de l'infrastructure dans Amazon Kendra

En tant que service géré, Amazon Kendra est protégé par la sécurité du réseau AWS mondial. Pour plus d'informations sur les services AWS de sécurité et sur la manière dont AWS l'infrastructure est protégée, consultez la section Sécurité du AWS cloud. Pour concevoir votre AWS environnement en utilisant les meilleures pratiques en matière de sécurité de l'infrastructure, consultez la section Protection de l'infrastructure dans le cadre AWS bien architecturé du pilier de sécurité.

Vous utilisez des appels d'API AWS publiés pour accéder à Amazon Kendra via le réseau. Les clients doivent prendre en charge les éléments suivants :

- Protocole TLS (Transport Layer Security). Nous exigeons TLS 1.2 et recommandons TLS 1.3.
- Ses suites de chiffrement PFS (Perfect Forward Secrecy) comme DHE (Ephemeral Diffie-Hellman) ou ECDHE (Elliptic Curve Ephemeral Diffie-Hellman). La plupart des systèmes modernes tels que Java 7 et les versions ultérieures prennent en charge ces modes.

En outre, les demandes doivent être signées à l’aide d’un ID de clé d’accès et d’une clé d’accès secrète associée à un principal IAM. Vous pouvez également utiliser AWS Security Token Service (AWS STS) pour générer des informations d’identification de sécurité temporaires et signer les demandes.

Analyse de configuration et de vulnérabilité dans AWS Identity and Access Management

AWS gère les tâches de sécurité de base telles que l'application de correctifs au système d'exploitation client (OS) et aux bases de données, la configuration du pare-feu et la reprise après sinistre. Ces procédures ont été vérifiées et certifiées par les tiers appropriés. Pour plus de détails, consultez les ressources suivantes :

- Modèle de responsabilité partagée
- AWS: Présentation des processus de sécurité (livre blanc)

Les ressources suivantes traitent également de la configuration et de l'analyse des vulnérabilités dans AWS Identity and Access Management (IAM) :

- Validation de conformité pour AWS Identity and Access Management
- Bonnes pratiques de sécurité et cas d'utilisation dans AWS Identity and Access Management.

Quotas pour Amazon Kendra

Régions prises en charge

Pour obtenir la liste des AWS régions où cette option Amazon Kendra est disponible, consultez la section Amazon Kendra Régions et points de terminaison dans le manuel Amazon Web Services General Reference.

> **Important**
>
> Les indices Amazon Kendra GenAI Enterprise Edition ne sont disponibles que dans les régions de l'Est des États-Unis (Virginie du Nord), de l'Ouest des États-Unis (Oregon), de l'Europe (Irlande) et de l'Asie-Pacifique (Sydney).

Quotas

Les quotas de service, également appelés limites, correspondent au nombre maximal de ressources de service pour votre AWS compte. Pour plus d'informations, veuillez consulter la rubrique Quotas du service Amazon Kendra dans les Références générales AWS .

> **Note**
>
> Les quotas pour les indices Amazon Kendra GenAI Enterprise Edition ne sont actuellement pas disponibles dans la console Service Quotas.

Quotas d'indice

Description	Par défaut	Edition	Ajustable
Nombre maximum d'indices par compte	10	GenAI Enterprise, développeur, entreprise	Oui
Quantité de texte extraite pour un index	3 Go	Developer	Non

Description	Par défaut	Edition	Ajustable
dans une seule unité (développeur). Vous ne pouvez pas ajouter d'unités supplémen taires pour extraire du texte pour l'édition Developer.			
Quantité de texte extraite pour un index dans une seule unité (Enterprise). Vous pouvez ajouter jusqu'à 100 unités supplémentaires pour extraire du texte pour l'édition Enterprise, ou simplement contacter le Support.	30 Go	Entreprise GenAI	Oui
Nombre maximum d'unités de capacité de stockage par indice GenAI Enterprise Edition	50	Enterprise	Oui

Quotas de connecteurs de sources de données

Description	Par défaut	Edition	Ajustable
Nombre maximal de connecteurs de source de données	5	Developer	Non

Description	Par défaut	Edition	Ajustable
par index (développeur)			
Nombre maximal de connecteurs de source de données par index (Enterprise)	50	Enterprise	Oui
Nombre maximal de connecteurs de source de données par index (GenAI Enterprise)	50	Entreprise GenAI	Oui
Nombre maximum d'heures d'utilisation de la synchronisation des connecteurs par GenAI Enterprise Index et par mois	500	Entreprise GenAI	Oui
Taille maximale d'un seul document ou d'un fichier brut pouvant être ingéré lors de l'utilisation d'un connecteur de source de données	50 Mo	GenAI Enterprise, développeur, entreprise	Oui

Description	Par défaut	Edition	Ajustable
Nombre maximal de préfixes S3 dans le fichier de configuration de la liste de contrôle d'accès inclus dans le connecteur de source de Amazon S3 données	100	Développeur, Enterprise	Non
Taille maximale du fichier de configuration de la liste de contrôle d'accès inclus dans le connecteur de source de Amazon S3 données	50 Mo	Développeur, Enterprise	Oui

FAQ sur les quotas

Description	Par défaut	Edition	Ajustable
Nombre maximum de FAQs par index	30	Développeur, Enterprise	Oui
Taille maximale d'une FAQ	5 Mo	Développeur, Enterprise	Oui
Nombre maximum de résultats renvoyés pour la FAQ	4	Développeur, Enterprise	Oui
Nombre maximum de caractères autorisés	300	Développeur, Enterprise	Non

Description	Par défaut	Edition	Ajustable
pour une question FAQ			
Nombre maximum de caractères dans une réponse à une FAQ	2000	Développeur, Enterprise	Non

Quotas du thésaurus

Description	Par défaut	Edition	Ajustable
Nombre maximum de thésaurus par index	1	Développeur, Enterprise	Non
Taille maximale d'un fichier de thésaurus	5 Mo	Développeur, Enterprise	Oui
Nombre maximum de règles de synonymes par thésaurus	10 000	Développeur, Enterprise	Oui
Nombre maximum de synonymes par terme dans tous les thésaurus d'un index	10	Développeur, Enterprise	Non

Amazon Kendra quotas d'expérience

Description	Par défaut	Edition	Ajustable
Nombre maximum d' Amazon Kendra expériences par indice	50	Développeur, Enterprise	Oui

Quotas de requêtes et de résultats de recherche

> **Note**
>
> Les requêtes par seconde sont partagées entre les API de récupération et de requête.

Description	Par défaut	Edition	Ajustable
Nombre de requêtes par seconde pour un index dans une seule unité (développeur). Vous ne pouvez pas ajouter d'unités supplémentaires pour les requêtes relatives à l'édition Developer.	0,05	Developer	Non
Nombre de requêtes par seconde pour un index dans une seule unité (Enterprise). Vous pouvez ajouter jusqu'à 100 unités supplémentaires pour toute demande concernant l'édition Enterprise, ou simplement contacter le Support.	0.1	Enterprise	Oui
Nombre maximum d'unités de capacité de requête par index GenAI Enterprise Edition	100	Entreprise GenAI	Oui

Description	Par défaut	Edition	Ajustable
Nombre maximum de caractères par texte de requête	1 000	GenAI Enterprise, développeur, entreprise	Oui
Nombre maximum de résultats de recherche par requête. La valeur par défaut est 100. Pour autoriser plus de 100 résultats, il suffit de contacter le Support.	100	Développeur, Enterprise	Oui
Nombre maximum de résultats de recherche par page	100	Développeur, Enterprise	Oui
Nombre maximal de mots symboliques par texte de requête avant troncature. La valeur par défaut est 30. Pour autoriser plus de 30 mots, contactez simplement le Support.	30	Développeur, Enterprise	Oui
Taille maximale de liste de groupes d'utilisateurs par attribut de requête	1 000	GenAI Enterprise, développeur, entreprise	Oui
Taille maximale de la liste de chaînes par attribut de requête	10	GenAI Enterprise, développeur, entreprise	Oui

Quotas de suggestions de requêtes

Description	Par défaut	Edition	Ajustable
Nombre maximum de suggestions de requêtes renvoyées par GetQuerySuggestionsappel	10	Développeur, Enterprise	Oui
Nombre maximum de champs/attributs pour les suggestions de requêtes par appel GetQuerySuggestions	10	Développeur, Enterprise	Oui
Nombre maximum de champs/attributs supplémentaires pour les suggestions de requêtes par appel GetQuerySuggestions	5	Développeur, Enterprise	Oui
Nombre maximum de listes de blocage par index	1	Développeur, Enterprise	Non
Taille maximale d'un fichier texte de liste de blocage	2 Mo	Développeur, Enterprise	Oui
Nombre maximum d'éléments (mots ou phrases) dans une liste de blocage	20 000	Développeur, Enterprise	Oui
Nombre maximal de suggestions de	1	Développeur, Enterprise	Oui

Description	Par défaut	Edition	Ajustable
requêtes corrigées orthographiquement à renvoyer lors d'un `Query` appel d'API.			

Quotas de documents

Description	Par défaut	Edition	Ajustable
Quantité de texte extraite pour un index dans une seule unité (développeur). Vous ne pouvez pas ajouter d'unités supplémentaires pour extraire du texte pour l'édition Developer.	3 Go	Developer	Non
Quantité de texte extraite pour un index dans une seule unité (Enterprise). Vous pouvez ajouter jusqu'à 100 unités supplémentaires pour extraire du texte pour l'édition Enterprise, ou simplement contacter le Support.	30 Go	Enterprise	Oui
Taille maximale d'un seul document ou d'un fichier brut lors de l'utilisation d'un	50 Mo	Développeur, Enterprise	Oui

Description	Par défaut	Edition	Ajustable
connecteur de source de données			
Taille maximale d'un seul document ou d'un fichier brut lors de l'utilisation de l'`BatchPutDocument` API	5 Mo	GenAI Enterprise, développeur, entreprise	Oui
Quantité maximale de texte extrait d'un seul document	5 Mo	Développeur, Enterprise	Non
Nombre maximum de champs/attributs personnalisés par index	500	Développeur, Enterprise	Non

Quotas de résultats de recherche en vedette

Description	Par défaut	Edition	Ajustable
Nombre maximum de documents en vedette par ensemble de résultats en vedette	4	Enterprise	Oui
Nombre maximum de textes de requête par ensemble de résultats en vedette	49	Enterprise	Non
Nombre maximum de caractères par texte	1 000	Enterprise	Oui

Description	Par défaut	Edition	Ajustable
de requête dans un ensemble de résultats en vedette			
Nombre maximum d'ensembles de résultats présentés par index	50	Enterprise	Oui

Rescore/ Quotas de résultats de recherche

Description	Par défaut	Edition	Ajustable
Nombre maximal de `Rescore` demandes par seconde pour un plan d'exécuti on de la nouvelle notation ou une seule unité de capacité. Vous pouvez ajouter jusqu'à 1 000 unités supplémentaires.	0,01	Enterprise	Non
Nombre maximum de plans d'exécution de la renotation par compte.	50	Enterprise	Oui
Nombre maximum de jetons `Title` pour un document dans une `Rescore` demande.	100	Enterprise	Non

Description	Par défaut	Edition	Ajustable
Nombre maximum de jetons `Body` pour un document dans une `Rescore` demande.	200	Enterprise	Non
Nombre maximum de documents par `Rescore` demande.	25	Enterprise	Non
Nombre maximum de documents par groupe dans une `Rescore` demande.	3	Enterprise	Non

Pour plus d'informations sur les quotas Amazon Kendra de service et pour demander une augmentation de quota, consultez Quotas de service.

Résolution des problèmes

Cette section peut vous aider à résoudre les problèmes courants que vous pourriez rencontrer lorsque vous travaillez avec Amazon Kendra.

Rubriques

- Dépannage des sources de données
- Résolution des problèmes liés aux résultats de recherche documentaire
- Dépannage de problèmes généraux

Dépannage des sources de données

Cette section peut vous aider à résoudre les problèmes courants liés à la configuration et à l'utilisation Amazon Kendra des connecteurs de source de données.

Mes documents n'ont pas été indexés

Lorsque vous synchronisez votre Amazon Kendra index avec une source de données, vous pouvez rencontrer des problèmes qui empêchent l'indexation des documents. L'indexation est un processus en deux étapes. Tout d'abord, la source de données est vérifiée pour détecter les documents nouveaux et mis à jour à indexer, et pour trouver les documents à supprimer de l'index. Ensuite, au niveau du document, chaque document est consulté et indexé.

Une erreur peut se produire lors de l'une ou l'autre de ces étapes. Les erreurs au niveau de la source de données sont signalées dans la console dans la section Historique des exécutions de synchronisation de la page de détails de la source de données. Le statut de la tâche de synchronisation peut être Réussi, Incomplet ou Échoué. Vous pouvez également voir le nombre de documents indexés et supprimés au cours de la tâche. Si le statut est Échoué, un message s'affiche dans la colonne Détails.

Les erreurs au niveau du document sont signalées dans Amazon CloudWatch Logs. Vous pouvez voir les erreurs à l'aide de la CloudWatch console.

Pour générer un rapport d'état de synchronisation de documents, voir Je souhaite générer un rapport d'état de synchronisation pour mes documents.

Ma tâche de synchronisation a échoué

Une tâche de synchronisation échoue généralement en cas d'erreur de configuration dans l'index ou dans la source de données. Dans la console, vous pouvez trouver le message d'erreur dans la section Historique des exécutions de la page de détails de la source de données, sous la colonne Détails. Les erreurs au niveau du document sont signalées dans Amazon CloudWatch Logs. Le message d'erreur fournit des informations sur ce qui s'est mal passé. Le problème est généralement que l'index ou la source de données ne disposent pas des IAM autorisations appropriées. Le message d'erreur décrit les autorisations manquantes. Voici certains des messages d'erreur que vous pouvez recevoir :

```
Failed to create log group for job. Please make sure that the IAM role
provided has sufficient permissions.
```

Si votre rôle d'index n'est pas autorisé à être utilisé CloudWatch, la source de données ne sera pas en mesure de créer un CloudWatch journal. Si cette erreur s'affiche, vous devez ajouter CloudWatch des autorisations au rôle d'index.

```
Failed to access Amazon S3 file prefix (bucket name) while trying to crawl
your metadata files. Please make sure the IAM role (ARN) provided has
sufficient permissions.
```

Lorsque vous utilisez une source de Amazon S3 données, vous Amazon Kendra devez être autorisé à accéder au compartiment contenant les documents. Vous devez ajouter l'autorisation Amazon Kendra de lire le bucket au IAM rôle de source de données.

```
The provided IAM role (ARN) could not be assumed. Please make sure Amazon
Kendra is a trusted entity that is allowed to assume the role.
```

Amazon Kendra a besoin d'une autorisation pour assumer les IAM rôles d'index et de source de données. Vous devez ajouter une politique de confiance aux rôles avec autorisation pour l'`sts:AssumeRole`action.

Pour les IAM politiques qui Amazon Kendra doivent indexer une source de données, consultez la section IAM rôles.

Pour générer un rapport d'état de synchronisation de documents, voir Je souhaite générer un rapport d'état de synchronisation pour mes documents.

Ma tâche de synchronisation est incomplète

Les tâches sont généralement incomplètes lorsqu'elles ont terminé le processus au niveau de la source de données, mais qu'elles comportent des erreurs au cours du processus au niveau du document. Lorsqu'une tâche est incomplète, il est possible que certains documents n'aient pas été correctement indexés. Dans le cas d'une source de Amazon S3 données, une tâche incomplète est généralement due à :

- Les métadonnées d'un ou de plusieurs documents n'étaient pas valides.
- Lorsque des documents sont soumis pour indexation mais qu'au moins un document n'a pas été soumis.
- Lorsque des documents sont soumis pour être supprimés de l'index mais qu'au moins un document n'a pas été soumis.

Pour résoudre les problèmes liés à une tâche de synchronisation incomplète, examinez d'abord vos CloudWatch journaux.

1. Dans la colonne des détails, choisissez Afficher les détails dans CloudWatch.
2. Consultez les messages d'erreur pour déterminer la cause de l'échec du document.

Pour générer un rapport d'état de synchronisation de documents, voir Je souhaite générer un rapport d'état de synchronisation pour mes documents.

Ma tâche de synchronisation a réussi mais aucun document n'est indexé

Parfois, une tâche de synchronisation d'index exécutée est marquée comme réussie, mais aucun document nouveau ou mis à jour n'est indexé comme prévu. Les raisons possibles sont les suivantes :

- Vérifiez la CloudWatch `DocumentsSubmittedForIndexingFailed` métrique pour voir si des documents n'ont pas pu être synchronisés. Consultez vos CloudWatch journaux pour plus de détails.
- Pour une source de Amazon S3 données, vous avez peut-être donné Amazon Kendra le mauvais nom de compartiment ou le mauvais préfixe. Assurez-vous que le bucket utilisé Amazon Kendra est celui qui contient les documents à indexer.

- Lorsque vous réindexez un document qui n'a pas pu être indexé dans une tâche précédente, vous Amazon Kendra ne l'indexez que si vous avez modifié le document ou le fichier de métadonnées associé.

Pour générer un rapport d'état de synchronisation de documents, voir Je souhaite générer un rapport d'état de synchronisation pour mes documents.

Je rencontre des problèmes de format de fichier lors de la synchronisation de ma source de données

Si vous rencontrez des problèmes de format de fichier lors de l'ajout de fichiers à votre source de données ou lors de la synchronisation de votre source de données, assurez-vous que vos types de documents sont Amazon Kendra pris en charge. Pour une liste des types de documents pris en charge par la Amazon Kendra section Types ou formats de documents.

Si vous utilisez l'`BatchPutDocument`API avec des fichiers texte brut, spécifiez-le `PLAIN_TEXT` comme type de contenu.

Je souhaite générer un rapport d'historique de synchronisation pour mes documents

Vous pouvez consulter un rapport d'historique des opérations de synchronisation au niveau du document dans le cadre CloudWatch de votre tâche de synchronisation des sources de données en sélectionnant Afficher le rapport. Un rapport d'historique des opérations de synchronisation contiendra des détails sur la progression et le statut de chaque document dans le cadre de la tâche de synchronisation. Il indique si un document a réussi, a échoué ou a été ignoré pendant les étapes d'analyse, de synchronisation et d'indexation. Vous trouverez également tous les messages d'erreur relatifs à des documents échoués ou ignorés. Si le rapport n'affiche pas les résultats d'une tâche de synchronisation en cours, il est possible que les journaux ne soient pas encore disponibles. Revenez plus tard au fur et à mesure que des données sont émises dans le rapport lorsque des événements se produisent pendant le processus de synchronisation.

Pour accéder à votre rapport sur l'historique des opérations de synchronisation, procédez comme suit :

1. Ouvrez la console Amazon Kendra à l'adresse. https://console.aws.amazon.com/kendra/
2. Dans le menu de navigation de gauche, sous Gestion des données, choisissez Sources de données, puis choisissez votre source de données.

3. Sur la page récapitulative de votre source de données, faites défiler l'écran vers le bas et sélectionnez l'onglet Historique de synchronisation.
4. Dans l'historique des exécutions de synchronisation, sélectionnez Actions.
5. Dans Actions, sélectionnez Afficher le rapport. Vous serez redirigé vers la CloudWatch console où vous pourrez accéder à votre rapport.

> **Note**
>
> L'historique des opérations de synchronisation enregistre si un document a été correctement indexé lors de l'ingestion, y compris les pièces jointes ACLs et les métadonnées, pour tous les connecteurs pris en charge par Amazon Kendra.

Si vous utilisez le connecteur Amazon S3 :

Outre l'affichage du rapport d'historique des opérations de synchronisation au niveau du document dans CloudWatch, vous pouvez générer des rapports d'historique de synchronisation pour chaque document de votre source de données Amazon S3 et le copier dans un compartiment. Amazon S3 Au cours de ce processus, vos données sont cryptées à l'aide de AWS KMS clés et vous seul pouvez les consulter. Le statut du document signalé peut être l'un des suivants : Echec, Terminé ou Réussite avec des erreurs. Avant de pouvoir générer des rapports d'état de synchronisation pour Amazon S3, vous devez effectuer les opérations suivantes :

- Ajoutez le principal Amazon Kendra de service suivant à votre politique Amazon S3 d'accès

```
{
    “Version”: “2012-10-17",
    “Statement”: [
        {
            “Sid”: “KendraS3Access”,
            “Effect”: “Allow”,
            “Principal”: {
                “Service”: “kendra.amazonaws.com”
            },
            “Action”: “s3:PutObject”,
            “Resource”: “arn:aws:s3:::your-manifest-bucket-name/*”
        }
    ]
}
```

- Créez un Amazon S3 bucket avec des autorisations d'accès pour Amazon Kendra

Si vous utilisez la console, pour générer un rapport d'historique de synchronisation pour Amazon S3, choisissez d'activer l'option Générer des rapports dans la section facultative de synchronisation des rapports d'historique sur la page de détails de la source de données. Entrez ensuite l'emplacement du Amazon S3 compartiment et choisissez parmi les options de configuration disponibles. Les rapports seront générés lors de la prochaine synchronisation une fois que vous aurez activé l'option Générer un rapport.

Si vous supprimez le Amazon S3 compartiment, vous perdrez vos données de journal et devrez en configurer un nouveau pour stocker les nouveaux rapports de synchronisation.

> **Note**
>
> Un rapport d'historique de synchronisation fournit uniquement des informations indiquant si un connecteur Amazon S3 a correctement exploré et ingéré des données.

Combien de temps prend la synchronisation d'une source de données ?

Si aucun document n'est mis à jour, le temps de synchronisation d'un Amazon Kendra index augmente de façon linéaire en fonction du nombre de documents. Par exemple, la synchronisation de 1 000 documents sans mise à jour prendrait environ cinq minutes et celle de 2 000 documents sans mise à jour prendrait environ 10 minutes. En cas de mise à jour des documents, le temps de synchronisation augmentera en fonction du nombre de documents mis à jour.

Quels sont les frais de synchronisation d'une source de données ?

Lorsque vous synchronisez votre index, il faut deux minutes pour le réchauffer et l'activer Amazon EC2 afin d'établir les connexions nécessaires. Vous n'êtes pas débité pendant ce processus. Votre compteur d'utilisation ne démarre qu'après le début de la tâche de synchronisation. Pour plus d'informations sur les Amazon Kendra tarifs, consultez la section Amazon Kendra tarification.

Je reçois une erreur Amazon EC2 d'autorisation

Si une erreur de fonctionnement Amazon EC2 non autorisée se produit lors de la synchronisation d'une source de données de cloud privé virtuel (VPC), il est probable que votre IAM rôle VPC ne

dispose pas des autorisations requises. Vérifiez que le IAM rôle que vous utilisez pour votre source de données dispose des autorisations associées. Pour plus d'informations, consultez la section IAM Rôle de cloud privé virtuel.

Je ne parviens pas à utiliser les liens de l'index de recherche pour ouvrir mes Amazon S3 objets

Votre Amazon Kendra index ne peut accéder qu'aux fichiers autorisés par une source de Amazon S3 données. Par exemple, Amazon Kendra impossible de modifier les Amazon S3 autorisations qui déterminent si un objet est censé être public ou chiffré. Amazon Kendra ne dispose pas non plus des autorisations par défaut pour créer ou renvoyer un lien signé pour Amazon S3 des objets. Si vous souhaitez activer les liens signés pour les Amazon S3 objets d'un Amazon Kendra index, deux options s'offrent à vous :

- Vous pouvez signer les résultats de votre requête d'index avec l'objet uri source avant de renvoyer le résultat sur la page de recherche. Pour une step-by-step présentation détaillée de ce processus, consultez la section Partage d'objets à l'aide d'une signature préalable URLs.
- Vous pouvez remplacer l'URI de la source des métadonnées de l' Amazon S3 objet et rendre votre service disponible via un réseau de diffusion de CloudFront contenu (CDN) connecté à un Amazon S3 bucket. Vous pouvez également utiliser un point de terminaison API Gateway proxy qui renvoie une URL présignée et redirige vers celle-ci.

Je reçois un message d'erreur AccessDenied lors de l'utilisation d'un fichier de certificat SSL

Si vous recevez un message d'erreur de refus d'accès lorsque vous utilisez un certificat SSL avec votre source de données, assurez-vous que votre IAM rôle est autorisé à accéder au fichier du certificat SSL à l'emplacement indiqué. Si le certificat est chiffré à l'aide d'une AWS KMS clé, votre IAM rôle doit également être autorisé à le déchiffrer à l'aide de cette AWS KMS clé. Pour plus d'informations, consultez Authentification et contrôle d'accès pour AWS KMS.

Je reçois une erreur d'autorisation lors de l'utilisation d'une source SharePoint de données

Si vous recevez une erreur d'autorisation lors de la synchronisation de votre index avec une source de SharePoint données, vérifiez qu'un rôle d'administrateur de site vous est attribué dans SharePoint.

Mon index n'explore pas les documents de ma source de données Confluence

Si votre Amazon Kendra index n'explore pas les documents de votre source de données Confluence pendant le processus de synchronisation, vérifiez que vous faites partie des groupes d'administrateurs de Confluence.

Résolution des problèmes liés aux résultats de recherche documentaire

Cette section peut vous aider à résoudre des problèmes dans les résultats Amazon Kendra de recherche.

Les résultats de ma recherche ne correspondent pas à ma requête

Si les résultats de votre recherche ne semblent pas pertinents, c'est peut-être pour les raisons suivantes :

- Les résultats obtenus en `LOW` toute confiance sont inclus dans les résultats. Vous pouvez filtrer les résultats `LOW` en toute confiance en utilisant le `ScoreAttributes` champ « s » pour exclure tout résultat dont la valeur QueryResultItemest égale à`LOW`. Amazon Kendra attribue à chaque résultat une valeur de compartiment de confiance de l'un ou l'autre `VERY_HIGHHIGH`, `MEDIUM` et`LOW`. Ces valeurs indiquent le niveau de confiance quant à la pertinence d'un résultat pour une requête. De plus, indépendamment des tranches de confiance, Amazon Kendra renvoie trois types de résultats dans l'ordre suivant : `ANSWER` (extrait de réponse suggéré), `QUESTION_ANSWER` (FAQ) et `DOCUMENT` (extrait de document). Par conséquent, il est possible qu'un `QUESTION_ANSWER` résultat de `LOW` confiance soit positionné au-dessus d'un `DOCUMENT` résultat de `VERY_HIGH` confiance. Cependant, il n'est pas toujours vrai que `LOW` la confiance `QUESTION_ANSWER` est un meilleur résultat que la `VERY_HIGH` confiance`DOCUMENT`.
- Certains champs ou attributs de métadonnées sont augmentés à une valeur très élevée, ce qui affecte le classement des résultats. Amazon Kendra effectue des recherches dans votre index à l'aide de plusieurs paramètres tels que le titre du document, le texte, la date et les champs de texte ou attributs personnalisés. Vous pouvez tester différentes valeurs d'amplification pour obtenir les meilleurs résultats pour toutes les requêtes. Vous pouvez également utiliser le réglage dynamique de la pertinence au niveau de la requête afin d'utiliser différentes valeurs d'amplification pour chaque requête.

- Vos utilisateurs utilisent des termes spécialisés lorsqu'ils demandent des informations et aucun synonyme personnalisé n'est configuré pour que votre index traite ces termes spécialisés. Pour plus de détails sur comment et quand utiliser des synonymes, voir Ajouter des synonymes personnalisés à un index.

Pourquoi est-ce que je ne vois que 100 résultats ?

Amazon Kendra renvoie le nombre total de documents pertinents. Les 100 premiers sont renvoyés par requête par défaut. Les résultats sont paginés. Vous pouvez l'utiliser `PageNumber` pour accéder à différentes pages.

Vous pouvez configurer Amazon Kendra pour renvoyer jusqu'à 1 000 documents ou résultats de recherche par requête, avec un maximum de 100 résultats par page. Pour renvoyer plus de 100 résultats, vous pouvez en faire la demande en contactant le Support des quotas. L'augmentation du nombre de résultats de recherche peut avoir un impact sur le temps de latence.

Pourquoi les documents que je m'attends à voir disparaissent-ils ?

Amazon Kendra prend en charge les listes de contrôle d'accès (ACLs) basées sur les utilisateurs et les groupes. Amazon Kendra ingère les politiques ACL via des connecteurs. Si un index ne configure pas d'ACL, seuls les documents correspondant au filtre d'attributs pour l'utilisateur et le groupe seront affichés. Si un filtre d'attribut d'utilisateur ou de groupe est fourni, les documents sans ACL ne seront pas affichés.

Si vous utilisez un contrôle d'accès basé sur des jetons, les documents sans politique ACL et les documents correspondant à l'utilisateur et aux groupes seront affichés.

Pourquoi est-ce que je vois des documents dotés d'une politique ACL ?

Si un index ne configure pas de politique de contrôle d'accès, le filtre peut fournir des utilisateurs et des groupes. Si aucun filtre d'utilisateur ou de groupe n'est appliqué, tous les documents associés seront renvoyés. Toute politique ACL sera ignorée.

Dépannage de problèmes généraux

Amazon Kendra utilise des CloudWatch métriques et des journaux pour fournir des informations sur la synchronisation de vos sources de données. Vous pouvez utiliser les métriques et les journaux

pour déterminer ce qui s'est mal passé lors d'une exécution de synchronisation et comment y remédier.

Pour un dépannage général, commencez par vos CloudWatch indicateurs.

- Vérifiez la `DocumentsCrawled` métrique pour connaître le nombre de documents vérifiés par votre source de données. Pour un Amazon S3 compartiment, si le nombre est inférieur à ce que vous espériez, vérifiez que votre source de données pointe vers le bon compartiment.
- Vérifiez la `DocumentsSkippedNoChange` métrique pour savoir combien de documents ont été ignorés car ils n'ont pas changé depuis la dernière synchronisation. Si le numéro ne correspond pas à ce que vous attendez, vérifiez que votre dépôt a été correctement mis à jour.
- Vérifiez la `DocumentsSkippedInvalidMetadata` métrique pour savoir combien de documents contenaient des métadonnées non valides. Consultez vos CloudWatch journaux pour voir les erreurs spécifiques qui se sont produites.
- Vérifiez la `DocumentsSubmittedForIndexingFailed` métrique pour savoir combien de documents ont été envoyés de la source de données à l'index mais n'ont pas pu être indexés. Par exemple, si vous utilisez un attribut de métadonnées dans une source de Amazon S3 données qui n'a pas été définie comme champ d'index personnalisé, le document ne sera pas indexé. Consultez vos CloudWatch journaux pour voir les erreurs spécifiques qui se sont produites.
- Vérifiez la `DocumentsSubmittedForDeletionFailed` métrique pour savoir combien de documents que la source de données a tenté de supprimer de l'index n'ont pas pu être supprimés de l'index. Consultez vos CloudWatch journaux pour voir les erreurs spécifiques qui se sont produites.

Vous pouvez consulter les CloudWatch journaux d'une exécution de synchronisation donnée pour obtenir des informations détaillées sur les erreurs survenues pendant l'exécution. Pour plus d'informations sur CloudWatch les journaux avec Amazon Kendra, consultez CloudWatch Logs.

Amazon Kendra Classement intelligent

Amazon Kendra Le classement intelligent utilise des fonctionnalités de recherche Amazon Kendra sémantique pour reclasser intelligemment les résultats d'un service de recherche.

Rubriques

- Amazon Kendra Classement intelligent pour l'autogestion OpenSearch
- Classement sémantique des résultats d'un service de recherche

Amazon Kendra Classement intelligent pour l'autogestion OpenSearch

Vous pouvez tirer Amazon Kendra parti des fonctionnalités de recherche sémantique du service de OpenSearchrecherche open source autogéré basé sur la licence Apache 2.0 pour améliorer les résultats de recherche. Le plugin Amazon Kendra Intelligent Ranking reclasse sémantiquement les résultats en utilisant OpenSearch. Amazon Kendra Pour ce faire, il comprend le sens et le contexte d'une requête de recherche à l'aide de champs spécifiques, tels que le corps ou le titre du document, à partir des résultats de OpenSearch recherche par défaut.

Prenons, par exemple, cette requête : « adresse principale du keynote ». Étant donné que le terme « adresse » a plusieurs significations, Amazon Kendra vous pouvez déduire le sens de la requête pour renvoyer des informations pertinentes conformes à la signification prévue. Dans ce contexte, il s'agit d'un discours liminaire de conférence. Un service de recherche plus simple risque de ne pas prendre en compte l'intention et de renvoyer des résultats pour une adresse postale sur Main Street, par exemple.

Le plugin Intelligent Ranking pour OpenSearch est disponible pour les versions OpenSearch (autogérées) 2.4.0 et ultérieures. Vous pouvez installer le plugin à l'aide d'un script Bash de démarrage rapide pour créer une nouvelle image Docker OpenSearch avec le plugin Intelligent Ranking inclus. Voir Configuration du plugin de recherche intelligent : il s'agit d'un exemple de configuration pour vous permettre d'être rapidement opérationnel.

Comment fonctionne le plugin de recherche intelligent

Le processus global du plugin Intelligent Ranking pour OpenSearch (autogéré) est le suivant :

1. Un OpenSearch utilisateur émet une requête et OpenSearch fournit une réponse à la requête ou une liste de documents pertinents pour la requête.
2. Le plugin Intelligent Ranking prend la réponse à la requête et extrait les informations des documents.
3. Le plugin Intelligent Ranking appelle l'API Rescore d' Amazon Kendra Intelligent Ranking.
4. L'`Rescore`API prend les informations extraites des documents et reclasse sémantiquement les résultats de recherche.
5. L'`Rescore`API renvoie les résultats de recherche reclassés au plugin. Le plugin réorganise les résultats de recherche dans la réponse de OpenSearch recherche pour refléter le nouveau classement sémantique.

Le plugin Intelligent Ranking reclasse les résultats en utilisant les champs « corps » et « titre ». Ces champs du plugin peuvent être mappés aux champs de votre OpenSearch index qui correspondent le mieux à la définition du corps et du titre d'un document. Par exemple, si votre index contient des chapitres d'un livre avec des champs tels que « chapter_heading » et « chapter_contents », vous pouvez associer le premier au « title » et le second à « body » pour obtenir les meilleurs résultats.

Configuration du plugin de recherche intelligent

Ce qui suit explique comment configurer rapidement OpenSearch (autogéré) le plugin Intelligent Ranking.

Configuration OpenSearch (autogérée) avec le plugin Intelligent Ranking (configuration rapide)

Si vous utilisez déjà l'image Docker`opensearch:2.4.0`, vous pouvez utiliser ce Dockerfile pour créer une nouvelle image de la version OpenSearch 2.4.0 avec le plugin Intelligent Ranking. Vous incluez un conteneur pour la nouvelle image dans votre fichier docker-compose.yml ou opensearch.yml. Vous devez également inclure l'ID du plan d'exécution de la renotation généré lors de la création d'un plan d'exécution de la renotation, ainsi que les informations relatives à votre région et à votre point de terminaison. Consultez l'étape 2 pour créer un plan d'exécution de la renotation.

Si vous avez déjà téléchargé une version de l'image `opensearch` Docker antérieure à la version 2.4.0, vous devez utiliser l'image Docker `opensearch:2.4.0` ou une version ultérieure et créer une nouvelle image avec le plugin Intelligent Ranking inclus.

1. Téléchargez et installez Docker Desktop pour votre système d'exploitation. Docker Desktop inclut Docker Compose et Docker Engine. Il est recommandé de vérifier si votre ordinateur répond à la configuration système requise mentionnée dans les détails d'installation de Docker.

   Vous pouvez également augmenter vos besoins en matière d'utilisation de la mémoire dans les paramètres de votre Docker Desktop. Vous êtes responsable des exigences d'utilisation de Docker en dehors des limites d'utilisation disponibles gratuitement pour les services Docker. Consultez la section Abonnements Docker.

   Vérifiez que l'état de Docker Desktop est « en cours d'exécution ».

2. Provisionnez Amazon Kendra un classement intelligent et vos exigences en matière de capacité. Une fois que vous avez configuré Amazon Kendra Intelligent Ranking, vous êtes facturé à l'heure en fonction des unités de capacité définies. Consultez le niveau gratuit et les informations tarifaires.

   Vous utilisez l'CreateRescoreExecutionPlanAPI pour approvisionner le`Rescore API`. Si vous n'avez pas besoin de plus d'unités de capacité que le nombre d'unités par défaut, n'ajoutez pas d'unités supplémentaires et donnez uniquement un nom à votre plan d'exécution de la nouvelle notation. Vous pouvez également mettre à jour vos exigences de capacité à l'aide de l'UpdateRescoreExecutionPlanAPI. Pour plus d'informations, consultez la section Classement sémantique des résultats d'un service de recherche.

   Vous pouvez éventuellement passer à l'étape 3 pour créer un plan d'exécution de la renotation par défaut lorsque vous exécutez le script Bash de démarrage rapide.

   Notez pour l'étape 4 l'ID du plan d'exécution de la renotation inclus dans la réponse.

   CLI

```
aws kendra-ranking create-rescore-execution-plan \
  --name MyRescoreExecutionPlan \
  --capacity-units '{"RescoreCapacityUnits":<integer number of additional
 capacity units>}'

Response:

{
    "Id": "<rescore execution plan ID>",
    "Arn": "arn:aws:kendra-ranking:<region>:<account-id>:rescore-execution-plan/
<rescore-execution-plan-id>"
```

```
}
```

Python

```
import boto3
from botocore.exceptions import ClientError
import pprint
import time

kendra_ranking = boto3.client("kendra-ranking")

print("Create a rescore execution plan.")

# Provide a name for the rescore execution plan
name = "MyRescoreExecutionPlan"
# Set your required additional capacity units
# Don't set capacity units if you don't require more than 1 unit given by
 default
capacity_units = 1

try:
    rescore_execution_plan_response =
 kendra_ranking.create_rescore_execution_plan(
        Name = name,
        CapacityUnits = {"RescoreCapacityUnits":capacity_units}
    )

    pprint.pprint(rescore_execution_plan_response)

    rescore_execution_plan_id = rescore_execution_plan_response["Id"]

    print("Wait for Amazon Kendra to create the rescore execution plan.")

    while True:
        # Get the details of the rescore execution plan, such as the status
        rescore_execution_plan_description =
 kendra_ranking.describe_rescore_execution_plan(
            Id = rescore_execution_plan_id
        )
        # When status is not CREATING quit.
        status = rescore_execution_plan_description["Status"]
        print(" Creating rescore execution plan. Status: "+status)
        time.sleep(60)
```

```
        if status != "CREATING":
            break

except ClientError as e:
        print("%s" % e)

print("Program ends.")
```

3. Téléchargez le script Bash de démarrage rapide correspondant GitHub à votre version de en OpenSearch sélectionnant la branche de version dans le menu déroulant des branches principales.

   Ce script utilise des images Docker OpenSearch et des OpenSearch tableaux de bord utilisant la version que vous avez sélectionnée dans le GitHub référentiel pour le script. Il télécharge un fichier zip pour le plugin Intelligent Ranking et génère une image Docker `Dockerfile` pour créer une nouvelle image Docker OpenSearch incluant le plugin. Il crée également un fichier docker-compose.yml qui inclut des conteneurs pour le plugin Intelligent Ranking et les tableaux OpenSearch de bord. OpenSearch Le script ajoute l'ID de votre plan d'exécution de la renotation, les informations de région et le point de terminaison (utilise la région) au fichier docker-compose.yml. Le script s'exécute ensuite `docker-compose up` pour démarrer les conteneurs OpenSearch avec Intelligent Ranking inclus et les OpenSearch tableaux de bord. Pour arrêter les récipients sans les retirer, courez`docker-compose stop`. Pour retirer les conteneurs, exécutez`docker-compose down`.

4. Ouvrez votre terminal et dans le répertoire du script Bash, exécutez la commande suivante.

   ```
   bash search_processing_kendra_quickstart.sh -p <execution-plan-id> -r <region>
   ```

   Lorsque vous exécutez cette commande, vous fournissez l'ID du plan d'exécution de la renotation que vous avez noté à l'étape 2 lorsque vous avez configuré Amazon Kendra Intelligent Ranking, ainsi que les informations de votre région. Vous pouvez éventuellement configurer le classement Amazon Kendra intelligent à l'aide de `--create-execution-plan` cette option. Cela crée un plan d'exécution de la nouvelle notation avec un nom et une capacité par défaut.

   Pour ne pas perdre votre index lorsque le conteneur éphémère par défaut est supprimé, vous pouvez conserver votre index d'une exécution à l'autre en fournissant le nom du volume de données à l'aide de l'`--volume-name`option. Si vous avez déjà créé un index, vous pouvez spécifier le volume dans votre fichier docker-compose.yml ou opensearch.yml. Pour conserver vos volumes intacts, ne les exécutez pas`docker-compose down -v`.

Le script Bash de démarrage rapide configure vos AWS informations d'identification dans le OpenSearch keystore pour vous connecter à Amazon Kendra Intelligent Ranking. Pour fournir vos AWS informations d'identification au script, utilisez l'`--profile`option permettant de spécifier le AWS profil. Si l'`--profile`option n'est pas spécifiée, le script Bash de démarrage rapide tente de lire les AWS informations d'identification (clé d'accès/clé secrète, jeton de session facultatif) à partir des variables d'environnement, puis à partir du profil par défaut. AWS Si l'`--profile`option n'est pas spécifiée et qu'aucune information d'identification n'est trouvée, le script ne transmettra pas d'informations d'identification au OpenSearch keystore. Si aucune information d'identification n'est spécifiée dans le OpenSearch keystore, le plugin vérifie toujours les informations d'identification dans la chaîne de fournisseurs d'informations d'identification par défaut, y compris les informations d'identification du Amazon ECS conteneur ou les informations d'identification du profil d'instance fournies via le service de Amazon EC2 métadonnées.

Assurez-vous d'avoir créé un IAM rôle doté des autorisations nécessaires pour invoquer Amazon Kendra Intelligent Ranking. Voici un exemple de IAM politique accordant l'autorisation d'utiliser l'`Rescore`API pour un plan d'exécution de la renotation spécifique :

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": "kendra-ranking:Rescore",
            "Resource": "arn:aws:kendra-ranking:${Region}:${Account}:rescore-
execution-plan/${RescoreExecutionPlanId}"
        }
    ]
}
```

Exemple de docker-compose.yml

Exemple de fichier docker-compose.yml utilisant la OpenSearch version 2.4.0 ou ultérieure avec le plugin Intelligent Ranking et Dashboards 2.4.0 ou version ultérieure. OpenSearch

```
version: '3'
networks:
```

```
  opensearch-net:
volumes:
  <volume-name>:
services:
  opensearch-node:
    image: <Docker image tag name of OpenSearch with Intelligent Ranking plugin>
    container_name: opensearch-node
    environment:
      - cluster.name=opensearch-cluster
      - node.name=opensearch-node
      - discovery.type=single-node
      - kendra_intelligent_ranking.service.endpoint=https://kendra-
ranking.<region>.api.aws
      - kendra_intelligent_ranking.service.region=<region>
      - kendra_intelligent_ranking.service.execution_plan_id=<rescore-execution-plan-
id>
    ulimits:
      memlock:
        soft: -1
        hard: -1
      nofile:
        soft: 65536
        hard: 65536
    ports:
      - 9200:9200
      - 9600:9600
    networks:
      - opensearch-net
  volumes:
    <docker-volume-name>:/usr/share/opensearch/data
  opensearch-dashboard:
   image: opensearchproject/opensearch-dashboards:<your-version>
   container_name: opensearch-dashboards
   ports:
     - 5601:5601
   environment:
     OPENSEARCH_HOSTS: '["https://opensearch-node:9200"]'
   networks:
     - opensearch-net
```

Exemple de Dockerfile et de création d'une image

Exemple d'utilisation de la OpenSearch version 2.4.0 ou ultérieure avec le plugin Intelligent Ranking.
`Dockerfile`

```
FROM opensearchproject/opensearch:<your-version>
RUN /usr/share/opensearch/bin/opensearch-plugin install --batch  https://github.com/
opensearch-project/search-processor/releases/download/<your-version>/search-
processor.zip
```

Création d'une image Docker pour OpenSearch avec le plugin Intelligent Ranking.

```
docker build --tag=<Docker image tag name of OpenSearch with Intelligent Ranking
 plugin>
```

Interaction avec le plugin de recherche intelligent

Une fois que vous avez configuré OpenSearch (autogéré) le plugin Intelligent Ranking, vous pouvez interagir avec le plugin à l'aide de commandes curl ou de bibliothèques OpenSearch clientes. Les informations d'identification par défaut pour accéder OpenSearch au plugin Intelligent Ranking sont le nom d'utilisateur « admin » et le mot de passe « admin ».

Pour appliquer les paramètres du plugin Intelligent Ranking à un OpenSearch index :

Curl

```
curl -XPUT "https://localhost:9200/<your-docs-index>/_settings" -u 'admin:admin' --
insecure -H 'Content-Type: application/json' -d'
{
  "index": {
    "plugin" : {
      "searchrelevance" : {
        "result_transformer" : {
          "kendra_intelligent_ranking": {
              "order": 1,
              "properties": {
                "title_field": "title_field_name_here",
                "body_field": "body_field_name_here"
              }
          }
        }
```

```
      }
    }
  }
}
'
```

Python

```
pip install opensearch-py

from opensearchpy import OpenSearch
host = 'localhost'
port = 9200
auth = ('admin', 'admin')

client = OpenSearch(
    hosts = [{'host': host, 'port': port}],
    http_compress = True, # enables gzip compression for request bodies
    http_auth = auth,
    # client_cert = client_cert_path,
    # client_key = client_key_path,
    use_ssl = True,
    verify_certs = False,
    ssl_assert_hostname = False,
    ssl_show_warn = False,
    ca_certs = ca_certs_path
)

setting_body = {
    "index": {
        "plugin" : {
            "searchrelevance" : {
                "result_transformer" : {
                    "kendra_intelligent_ranking": {
                            "order": 1,
                            "properties": {
                                "title_field": "title_field_name_here",
                                "body_field": "body_field_name_here"
                            }
                    }
                }
            }
        }
```

```
    }
}

response = client.indices.put_settings(index_name, body=setting_body)
```

Vous devez inclure le nom du champ de texte principal que vous souhaitez utiliser pour le reclassement, tel qu'un corps de document ou un champ de contenu de document. Vous pouvez également inclure d'autres champs de texte, tels que le titre du document ou le résumé du document.

Vous pouvez désormais émettre n'importe quelle requête et les résultats sont classés à l'aide du plugin Intelligent Ranking.

Curl

```
curl -XGET "https://localhost:9200/<your-docs-index>/_search?pretty" -u
 'admin:admin' --insecure -H 'Content-Type: application/json' -d'
{
  "query" : {
    "match" : {
      "body_field_name_here": "intelligent systems"
    }
  }
}
'
```

Python

```
from opensearchpy import OpenSearch
host = 'localhost'
port = 9200
auth = ('admin', 'admin')

client = OpenSearch(
    hosts = [{'host': host, 'port': port}],
    http_compress = True, # enables gzip compression for request bodies
    http_auth = auth,
    # client_cert = client_cert_path,
    # client_key = client_key_path,
    use_ssl = True,
    verify_certs = False,
    ssl_assert_hostname = False,
```

```
    ssl_show_warn = False,
    ca_certs = ca_certs_path
)

query = {
  'size': 10,
  "query" : {
    "match" : {
      "body_field_name_here": "intelligent systems"
    }
  }
}

response = client.search(
    body = query,
    index = index_name
)

print('\nSearch results:')
print(response)
```

Pour supprimer les paramètres du plugin Intelligent Ranking pour un OpenSearch index, procédez comme suit :

Curl

```
curl -XPUT "http://localhost:9200/<your-docs-index>/_settings" -H 'Content-Type:
 application/json' -d'
{
  "index": {
    "plugin": {
      "searchrelevance": {
        "result_transformer": {
          "kendra_intelligent_ranking.*": null
        }
      }
    }
  }
}
'
```

Python

```
from opensearchpy import OpenSearch
host = 'localhost'
port = 9200
auth = ('admin', 'admin')

client = OpenSearch(
    hosts = [{'host': host, 'port': port}],
    http_compress = True, # enables gzip compression for request bodies
    http_auth = auth,
    # client_cert = client_cert_path,
    # client_key = client_key_path,
    use_ssl = True,
    verify_certs = False,
    ssl_assert_hostname = False,
    ssl_show_warn = False,
    ca_certs = ca_certs_path
)

setting_body = {
  "index": {
    "plugin": {
      "searchrelevance": {
        "result_transformer": {
          "kendra_intelligent_ranking.*": null
        }
      }
    }
  }
}

response = client.indices.put_settings(index_name, body=setting_body)
```

Pour tester le plugin Intelligent Ranking sur une certaine requête ou pour tester sur certains champs du corps et du titre :

Curl

```
curl -XGET "https://localhost:9200/<your-docs-index>/_search?pretty" -u
 'admin:admin' --insecure -H 'Content-Type: application/json' -d'
{
```

```
  "query": {
    "multi-match": {
      "query": "intelligent systems",
      "fields": ["body_field_name_here", "title_field_name_here"]
    }
  },
  "size": 25,
  "ext": {
    "search_configuration": {
      "result_transformer": {
        "kendra_intelligent_ranking": {
          "order": 1,
          "properties": {
            "title_field": "title_field_name_here",
            "body_field": "body_field_name_here"
          }
        }
      }
    }
  }
}
'
```

Python

```
from opensearchpy import OpenSearch
host = 'localhost'
port = 9200
auth = ('admin', 'admin')

client = OpenSearch(
    hosts = [{'host': host, 'port': port}],
    http_compress = True, # enables gzip compression for request bodies
    http_auth = auth,
    # client_cert = client_cert_path,
    # client_key = client_key_path,
    use_ssl = True,
    verify_certs = False,
    ssl_assert_hostname = False,
    ssl_show_warn = False,
    ca_certs = ca_certs_path
)
```

```
# Index settings null for kendra_intelligent_ranking

query = {
  "query": {
    "multi_match": {
      "query": "intelligent systems",
      "fields": ["body_field_name_here", "title_field_name_here"]
    }
  },
  "size": 25,
  "ext": {
    "search_configuration": {
      "result_transformer": {
        "kendra_intelligent_ranking": {
          "order": 1,
          "properties": {
            "title_field": "title_field_name_here",
            "body_field": "body_field_name_here"
          }
        }
      }
    }
  }
}

response = client.search(
    body = query,
    index = index_name
)

print('\nSearch results:')
print(response)
```

Comparaison des OpenSearch résultats avec les Amazon Kendra résultats

Vous pouvez comparer les résultats classés side-by-side OpenSearch (autogérés) aux Amazon Kendra résultats reclassés. OpenSearch Les versions 2.4.0 et ultérieures de Dashboards offrent des side-by-side résultats qui vous permettent de comparer le OpenSearch classement des documents à la manière Amazon Kendra dont le plugin classe les documents pour une requête de recherche.

Avant de pouvoir comparer les résultats OpenSearch classés aux résultats Amazon Kendra reclassés, assurez-vous que vos OpenSearch tableaux de bord sont sauvegardés par un

OpenSearch serveur doté du plugin Intelligent Ranking. Vous pouvez le configurer à l'aide de Docker et d'un script Bash de démarrage rapide. Consultez Configuration du plugin de recherche intelligent.

Ce qui suit explique comment comparer OpenSearch et Amazon Kendra rechercher des résultats dans les OpenSearch tableaux de bord. Pour de plus amples d'informations, consultez la documentationOpenSearch.

Comparaison des résultats de recherche dans les OpenSearch tableaux de bord

1. Ouvrez http://localhost:5601 et connectez-vous à OpenSearch Dashboards. Les informations d'identification par défaut sont le nom d'utilisateur « admin » et le mot de passe « admin ».
2. Sélectionnez Search Relevance dans les OpenSearch plugins du menu de navigation.
3. Entrez le texte de recherche dans la barre de recherche.
4. Sélectionnez votre index pour la requête 1 et entrez une requête dans le OpenSearch Query DSL. Vous pouvez utiliser la `%SearchText%` variable pour faire référence au texte de recherche que vous avez saisi dans la barre de recherche. Pour un exemple de cette requête, consultez OpenSearch la documentation. Les résultats renvoyés pour cette requête sont les OpenSearch résultats obtenus sans utiliser le plugin Intelligent Ranking.
5. Sélectionnez le même index pour la requête 2 et entrez la même requête dans le OpenSearch Query DSL. En outre, incluez l'extension `kendra_intelligent_ranking` et spécifiez l'extension obligatoire sur laquelle le classement doit `body_field` être effectué. Vous pouvez également spécifier le champ de titre, mais le champ de corps est obligatoire. Pour un exemple de cette requête, consultez OpenSearch la documentation. Les résultats renvoyés pour cette requête sont les résultats Amazon Kendra reclassés à l'aide du plugin Intelligent Ranking. Le plugin classe jusqu'à 25 résultats.
6. Sélectionnez Rechercher pour retourner et comparer les résultats.

Classement sémantique des résultats d'un service de recherche

Amazon Kendra Le classement intelligent utilise les fonctionnalités Amazon Kendra de recherche sémantique pour reclasser les résultats d'un service de recherche. Pour ce faire, il prend en compte le contexte de la requête de recherche, ainsi que toutes les informations disponibles dans les documents du service de recherche. Amazon Kendra Le classement intelligent peut améliorer la correspondance simple des mots clés.

L'CreateRescoreExecutionPlanAPI crée une ressource de classement Amazon Kendra intelligent utilisée pour le provisionnement de l'API Rescore. L'`Rescore`API reclasse les résultats de recherche provenant d'un service de recherche tel que OpenSearch (autogéré).

Lorsque vous appelez`CreateRescoreExecutionPlan`, vous définissez les unités de capacité requises pour reclasser les résultats d'un service de recherche. Si vous n'avez pas besoin d'unités de capacité supplémentaires au-delà de la valeur par défaut d'une unité, ne modifiez pas la valeur par défaut. Donnez uniquement un nom à votre plan d'exécution de la renotation. Vous pouvez configurer jusqu'à 1 000 unités supplémentaires. Pour plus d'informations sur ce qui est inclus dans une unité de capacité unique, voir Réglage de la capacité. Une fois que vous avez configuré Amazon Kendra Intelligent Ranking, vous êtes facturé à l'heure en fonction des unités de capacité que vous avez définies. Consultez le niveau gratuit et les informations tarifaires.

Un identifiant de plan d'exécution de la renotation est généré et renvoyé dans la réponse lorsque vous appelez`CreateRescoreExecutionPlan`. L'`Rescore`API utilise l'ID du plan d'exécution Rescore pour reclasser les résultats d'un service de recherche en fonction de la capacité que vous avez définie. Vous incluez l'ID du plan d'exécution de la renotation dans les fichiers de configuration de votre service de recherche. Par exemple, si vous utilisez OpenSearch (autogéré), vous incluez l'ID du plan d'exécution de la renotation dans votre fichier docker-compose.yml ou opensearch.yml. Voir Classement intelligent des résultats (en libre-service). OpenSearch

Un Amazon Resource Name (ARN) est également généré dans la réponse lorsque vous appelez`CreateRescoreExecutionPlan`. Vous pouvez utiliser cet ARN pour créer une politique d'autorisation dans AWS Identity and Access Management (IAM) afin de restreindre l'accès des utilisateurs à un ARN spécifique pour un plan d'exécution de la renotation spécifique. Pour un exemple de IAM politique autorisant l'utilisation de l'`Rescore`API pour un plan d'exécution de la renotation spécifique, voir Amazon Kendra Intelligent Ranking pour l' OpenSearchautogestion.

Voici un exemple de création d'un plan d'exécution de la nouvelle notation avec des unités de capacité définies sur 1.

CLI

```
aws kendra-ranking create-rescore-execution-plan \
  --name MyRescoreExecutionPlan \
  --capacity-units '{"RescoreCapacityUnits":1}'

Response:

{
```

```
    "Id": "<rescore execution plan ID>",
    "Arn": "arn:aws:kendra-ranking:<region>:<account-id>:rescore-execution-plan/
<rescore-execution-plan-id>"
}
```

Python

```
import boto3
from botocore.exceptions import ClientError
import pprint
import time

kendra_ranking = boto3.client("kendra-ranking")

print("Create a rescore execution plan.")

# Provide a name for the rescore execution plan
name = "MyRescoreExecutionPlan"
# Set your required additional capacity units
# Don't set capacity units if you don't require more than 1 unit given by default
capacity_units = 1

try:
    rescore_execution_plan_response = kendra_ranking.create_rescore_execution_plan(
        Name = name,
        CapacityUnits = {"RescoreCapacityUnits":capacity_units}
    )

    pprint.pprint(rescore_execution_plan_response)

    rescore_execution_plan_id = rescore_execution_plan_response["Id"]

    print("Wait for Amazon Kendra to create the rescore execution plan.")

    while True:
        # Get the details of the rescore execution plan, such as the status
        rescore_execution_plan_description =
 kendra_ranking.describe_rescore_execution_plan(
            Id = rescore_execution_plan_id
        )
        # When status is not CREATING quit.
        status = rescore_execution_plan_description["Status"]
        print(" Creating rescore execution plan. Status: "+status)
```

```
        time.sleep(60)
        if status != "CREATING":
            break

except ClientError as e:
        print("%s" % e)

print("Program ends.")
```

Java

```
import java.util.concurrent.TimeUnit;

import software.amazon.awssdk.services.kendraranking.KendraRankingClient;
import
 software.amazon.awssdk.services.kendraranking.model.CapacityUnitsConfiguration;
import
 software.amazon.awssdk.services.kendraranking.model.CreateRescoreExecutionPlanRequest;
import
 software.amazon.awssdk.services.kendraranking.model.CreateRescoreExecutionPlanResponse;
import
 software.amazon.awssdk.services.kendraranking.model.DescribeRescoreExecutionPlanRequest;
import
 software.amazon.awssdk.services.kendraranking.model.DescribeRescoreExecutionPlanResponse;
import
 software.amazon.awssdk.services.kendraranking.model.RescoreExecutionPlanStatus;

public class CreateRescoreExecutionPlanExample {

  public static void main(String[] args) throws InterruptedException {

    String rescoreExecutionPlanName = "MyRescoreExecutionPlan";
    int capacityUnits = 1;

    KendraRankingClient kendraRankingClient = KendraRankingClient.builder().build();

    System.out.println(String.format("Creating a rescore execution plan named %s",
 rescoreExecutionPlanName));

    CreateRescoreExecutionPlanResponse createResponse =
 kendraRankingClient.createRescoreExecutionPlan(
        CreateRescoreExecutionPlanRequest.builder()
            .name(rescoreExecutionPlanName)
```

```
            .capacityUnits(
                CapacityUnitsConfiguration.builder()
                    .rescoreCapacityUnits(capacityUnits)
                    .build()
            )
            .build()
    );

    String rescoreExecutionPlanId = createResponse.id();
    System.out.println(String.format("Waiting for rescore execution plan with id %s
 to finish creating.", rescoreExecutionPlanId));
    while (true) {
      DescribeRescoreExecutionPlanResponse describeResponse =
 kendraRankingClient.describeRescoreExecutionPlan(
          DescribeRescoreExecutionPlanRequest.builder()
              .id(rescoreExecutionPlanId)
              .build()
      );
      RescoreExecutionPlanStatus rescoreExecutionPlanStatus =
 describeResponse.status();
      if (rescoreExecutionPlanStatus != RescoreExecutionPlanStatus.CREATING) {
        break;
      }
      TimeUnit.SECONDS.sleep(60);
    }

    System.out.println("Rescore execution plan creation is complete.");
  }
}
```

Voici un exemple de mise à jour d'un plan d'exécution de la renotation pour définir les unités de capacité sur 2.

CLI

```
aws kendra-ranking update-rescore-execution-plan \
  --id <rescore execution plan ID> \
  --capacity-units '{"RescoreCapacityUnits":2}'
```

Python

```
import boto3
```

```
from botocore.exceptions import ClientError
import pprint
import time

kendra_ranking = boto3.client("kendra-ranking")

print("Update a rescore execution plan.")

# Provide the ID of the rescore execution plan
id = <rescore execution plan ID>
# Re-set your required additional capacity units
capacity_units = 2

try:
    kendra_ranking.update_rescore_execution_plan(
        Id = id,
        CapacityUnits = {"RescoreCapacityUnits":capacity_units}
    )

    print("Wait for Amazon Kendra to update the rescore execution plan.")

    while True:
        # Get the details of the rescore execution plan, such as the status
        rescore_execution_plan_description =
 kendra_ranking.describe_rescore_execution_plan(
            Id = id
        )
        # When status is not UPDATING quit.
        status = rescore_execution_plan_description["Status"]
        print(" Updating rescore execution plan. Status: "+status)
        time.sleep(60)
        if status != "UPDATING":
            break

except ClientError as e:
        print("%s" % e)

print("Program ends.")
```

Java

```
import java.util.concurrent.TimeUnit;
```

```
import software.amazon.awssdk.services.kendraranking.KendraRankingClient;
import
 software.amazon.awssdk.services.kendraranking.model.CapacityUnitsConfiguration;
import
 software.amazon.awssdk.services.kendraranking.model.DescribeRescoreExecutionPlanRequest;
import
 software.amazon.awssdk.services.kendraranking.model.DescribeRescoreExecutionPlanResponse;
import
 software.amazon.awssdk.services.kendraranking.model.RescoreExecutionPlanStatus;
import
 software.amazon.awssdk.services.kendraranking.model.UpdateRescoreExecutionPlanRequest;
import
 software.amazon.awssdk.services.kendraranking.model.UpdateRescoreExecutionPlanResponse;

public class UpdateRescoreExecutionPlanExample {

  public static void main(String[] args) throws InterruptedException {

    String rescoreExecutionPlanId = <rescore execution plan ID>;
    int newCapacityUnits = 2;

    KendraRankingClient kendraRankingClient = KendraRankingClient.builder().build();

    System.out.println(String.format("Updating a rescore execution plan named %s",
 rescoreExecutionPlanId));

    UpdateRescoreExecutionPlanResponse updateResponse =
 kendraRankingClient.updateRescoreExecutionPlan(
        UpdateRescoreExecutionPlanRequest.builder()
            .id(rescoreExecutionPlanId)
            .capacityUnits(
                CapacityUnitsConfiguration.builder()
                    .rescoreCapacityUnits(newCapacityUnits)
                    .build()
            )
            .build()
    );

    System.out.println(String.format("Waiting for rescore execution plan with id %s
 to finish updating.", rescoreExecutionPlanId));
    while (true) {
      DescribeRescoreExecutionPlanResponse describeResponse =
 kendraRankingClient.describeRescoreExecutionPlan(
          DescribeRescoreExecutionPlanRequest.builder()
```

```
                .id(rescoreExecutionPlanId)
                .build()
      );
      RescoreExecutionPlanStatus rescoreExecutionPlanStatus =
 describeResponse.status();
      if (rescoreExecutionPlanStatus != RescoreExecutionPlanStatus.UPDATING) {
        break;
      }
      TimeUnit.SECONDS.sleep(60);
    }

    System.out.println("Rescore execution plan update is complete.");
  }
}
```

Voici un exemple d'utilisation de l'`Rescore`API.

CLI

```
aws kendra-ranking rescore \
  --rescore-execution-plan-id <rescore execution plan ID> \
  --search-query "intelligent systems" \
  --documents "[{\"Id\": \"DocId1\",\"Title\": \"Smart systems\", \"Body\":
 \"intelligent systems in everyday life\",\"OriginalScore\": 2.0}, {\"Id\":
 \"DocId2\",\"Title\": \"Smarter systems\", \"Body\": \"living with intelligent
 systems\",\"OriginalScore\": 1.0}]"
```

Python

```
import boto3
from botocore.exceptions import ClientError
import pprint

kendra_ranking = boto3.client("kendra-ranking")

print("Use the Rescore API.")

# Provide the ID of the rescore execution plan
id = <rescore execution plan ID>
# The search query from the search service
query = "intelligent systems"
# The list of documents for Intelligent Ranking to rescore
```

```
document_list = [
    {"Id": "DocId1", "Title": "Smart systems", "Body": "intelligent systems in
 everyday life", "OriginalScore": 2.0},
    {"Id": "DocId2", "Title": "Smarter systems", "Body": "living with intelligent
 systems", "OriginalScore": 1.0}
]

try:
    rescore_response = kendra_ranking.rescore(
        rescore_execution_plan_id = id,
        search_query = query,
        documents = document_list
    )

    print(rescore_response["RescoreId"])
    print(rescore_resposne["ResultItems"])

except ClientError as e:
        print("%s" % e)

print("Program ends.")
```

Java

```
import java.util.ArrayList;
import java.util.List;

import software.amazon.awssdk.services.kendraranking.KendraRankingClient;
import software.amazon.awssdk.services.kendraranking.model.RescoreRequest;
import software.amazon.awssdk.services.kendraranking.model.RescoreResponse;
import software.amazon.awssdk.services.kendraranking.model.Document;

public class RescoreExample {

  public static void main(String[] args) {

    String rescoreExecutionPlanId = <rescore execution plan ID>;
    String query = "intelligent systems";

    List<Document> documentList = new ArrayList<>();
    documentList.add(
        Document.builder()
            .id("DocId1")
```

```
            .originalScore(2.0F)
            .body("intelligent systems in everyday life")
            .title("Smart systems")
            .build()
    );
    documentList.add(
        Document.builder()
            .id("DocId2")
            .originalScore(1.0F)
            .body("living with intelligent systems")
            .title("Smarter systems")
            .build()
    );

    KendraRankingClient kendraRankingClient = KendraRankingClient.builder().build();

    RescoreResponse rescoreResponse = kendraRankingClient.rescore(
        RescoreRequest.builder()
            .rescoreExecutionPlanId(rescoreExecutionPlanId)
            .searchQuery(query)
            .documents(documentList)
            .build()
    );

    System.out.println(rescoreResponse.rescoreId());
    System.out.println(rescoreResponse.resultItems());
  }
}
```

Référence d’API

La documentation de référence de l'API est désormais un guide distinct.

Historique du document pour Amazon Kendra

- Dernière mise à jour de la documentation : 4 décembre 2024

Le tableau suivant décrit les modifications importantes apportées à chaque version de Amazon Kendra. Pour recevoir les notifications sur les mises à jour de cette documentation, vous pouvez vous abonner au Flux RSS.

Modification	Description	Date
Nouvelle fonction	Amazon Kendra prend désormais en charge un nouveau type d'index : l'index Amazon Kendra GenAI Enterprise Edition. Pour plus d'informations, consultez l'index Amazon Kendra GenAI Enterprise Edition.	4 décembre 2024
Nouvelle fonction	Amazon Kendra prend désormais en charge une version mise à jour du connecteur de source de GitHub données. Pour de plus amples informations, veuillez consulter GitHub.	27 février 2024
Nouvelle fonction	Amazon Kendra prend désormais en charge une version mise à jour du connecteur de source de Amazon FSx données. Pour plus d'informations, consultez Amazon FSx (Windows) et Amazon FSx (NetAppONTAP).	8 février 2024

Nouvelle fonction	Amazon Kendra prend désormais en charge une version mise à jour du connecteur de source de données Slack. Pour plus d'informations, consultez Slack.	11 janvier 2024
Nouvelle fonction	Amazon Kendra prend désormais en charge la réduction et l'extension de vos résultats de recherche. Pour plus d'informations, consultez la section Réduction/extension des résultats de recherche.	19 octobre 2023
Nouvelle fonction	Amazon Kendra supporte désormais un connecteur de source de données Aurora (MySQL). Pour plus d'informations, consultez Aurora (MySQL).	28 septembre 2023
Nouvelle fonction	Amazon Kendra prend désormais en charge un Aurora connecteur de source de données (PostgreSQL). Pour plus d'informations, consultez Aurora (PostgreSQL).	28 septembre 2023
Nouvelle fonction	Amazon Kendra supporte désormais un connecteur de source de données Amazon RDS (MySQL). Pour plus d'informations, consultez Amazon RDS (MySQL).	28 septembre 2023

Nouvelle fonction	Amazon Kendra prend désormais en charge un connecteur de source de données Amazon RDS (Microsoft SQL Server). Pour plus d'informations, consultez Amazon RDS (Microsoft SQL Server).	28 septembre 2023
Nouvelle fonction	Amazon Kendra prend désormais en charge un connecteur de source de données Amazon RDS (Oracle). Pour plus d'informations, consultez Amazon RDS (Oracle).	28 septembre 2023
Nouvelle fonction	Amazon Kendra prend désormais en charge un Amazon RDS connecteur de source de données (PostgreSQL). Pour plus d'informations, consultez Amazon RDS (PostgreSQL).	28 septembre 2023
Nouvelle fonction	Amazon Kendra prend désormais en charge un connecteur DB2 de source de données IBM. Pour plus d'informations, consultez IBM DB2.	28 septembre 2023

Nouvelle fonction	Amazon Kendra prend désormais en charge un connecteur de source de données Microsoft SQL Server. Pour plus d'informations, consultez Microsoft SQL Server.	28 septembre 2023
Nouvelle fonction	Amazon Kendra supporte désormais un connecteur de source de données MySQL. Pour plus d'informations, consultez MySQL.	28 septembre 2023
Nouvelle fonction	Amazon Kendra prend désormais en charge un connecteur de source de données Oracle Database. Pour plus d'informations, consultez Oracle Database.	28 septembre 2023
Nouvelle fonction	Amazon Kendra prend désormais en charge un connecteur de source de données PostgreSQL. Pour plus d'informations, consultez PostgreSQL.	28 septembre 2023
Nouvelle fonction	Amazon Kendra fournit désormais un connecteur de source de données pour Drupal. Pour plus d'informations, consultez Drupal.	6 septembre 2023

Nouvelle fonction	Récupérez des passages sémantiquement pertinents à l'aide de l'API Amazon Kendra Retrieve pour les systèmes de génération augmentée (RAG).	22 juin 2023
Nouvelle fonction	Amazon Kendra prend désormais en charge une version mise à jour du connecteur de source de données Amazon Kendra Web Crawler. Pour plus d'informations, consultez Amazon Kendra Web Crawler v2.0.	21 juin 2023
Expansion de région	Amazon Kendra est désormais disponible en Europe (Londres) (eu-west-2).	5 juin 2023
Nouvelle fonction	Amazon Kendra prend désormais en charge une version mise à jour du connecteur de source de données Alfresco. Pour plus d'informations, consultez Alfresco.	16 mai 2023
Nouvelle fonction	Amazon Kendra fournit désormais un connecteur de source de données pour Adobe Experience Manager. Pour plus d'informations, consultez Adobe Experience Manager.	11 mai 2023

Nouvelle fonction	Amazon Kendra prend désormais en charge la configuration des champs/attributs du document lorsque vous appelez. GetQuerySuggestions Vous pouvez désormais baser les suggestions de requêtes sur le contenu des champs du document. Pour plus d'informations, consultez la section Suggestions de requêtes.	2 mai 2023
Nouvelle fonction	Amazon Kendra fournit désormais un connecteur de source de données pour Gmail. Pour plus d'informations, consultez Gmail.	13 avril 2023
Nouvelle fonction	Amazon Kendra prend désormais en charge une version mise à jour du connecteur de source de OneDrive données Microsoft. Pour plus d'informations, consultez Microsoft OneDrive v2.0.	3 avril 2023
Nouvelle fonction	Améliorez la visibilité des nouveaux documents ou faites la promotion de certains documents lorsque vos utilisateurs saisissent certaines requêtes à l'aide des résultats en vedette.	30 mars 2023

Nouvelle fonction	Amazon Kendra prend désormais en charge un connecteur de source de données mis à jour pour Microsoft SharePoint. Pour plus d'informations, consultez Microsoft SharePoint.	2 mars 2023
Nouvelle fonction	Amazon Kendra prend désormais en charge une version mise à jour du connecteur de source de données Confluence. Pour plus d'informations, consultez Confluence.	1er mars 2023
Expansion de région	Amazon Kendra est désormais disponible en Asie-Pacifique (Tokyo) (ap-northeast-1).	7 février 2023
Nouvelle fonction	Amazon Kendra fournit désormais un connecteur de source de données pour Microsoft Exchange. Pour plus d'informations, consultez Microsoft Exchange.	12 janvier 2023
Nouvelle fonction	Amazon Kendra fournit désormais un connecteur de source de données pour Microsoft Yammer. Pour plus d'informations, consultez Microsoft Yammer.	12 janvier 2023

Nouvelle fonction	Amazon Kendra prend désormais en charge l'indexation des types de documents RTF, XML, XSLT, MS_EXCEL, CSV, JSON et MD. Pour plus d'informations, consultez la section Types de documents.	11 janvier 2023
Nouvelle fonction	Amazon Kendra prend désormais en charge une version mise à jour du connecteur de source de Amazon S3 données. Pour de plus amples informations, veuillez consulter Amazon S3.	10 janvier 2023
Nouvelle fonction	OpenSearchLes résultats de recherche (autogérés) peuvent être classés sémantiquement à l'aide du classement Amazon Kendra intelligent.	9 janvier 2023
Nouvelle fonction	Amazon Kendra fournit désormais un connecteur de source de données pour Microsoft Teams. Pour plus d'informations, consultez Microsoft Teams.	5 janvier 2023
Nouvelle fonction	Amazon Kendra dispose d'un connecteur de source de données mis à jour pour Google Drive. Pour plus d'informations, consultez Google Drive.	5 janvier 2023

Nouvelle fonction	Amazon Kendra dispose d'un connecteur de source de données mis à jour pour ServiceNow. Pour de plus amples informations, veuillez consulter ServiceNow.	21 décembre 2022
Nouvelle fonction	Amazon Kendra dispose d'un connecteur de source de données mis à jour pour Salesforce. Pour plus d'informations, consultez Salesforce.	21 décembre 2022
Expansion de région	Amazon Kendra est désormais disponible en Asie-Pacifique (Mumbai) (ap-south-1).	14 décembre 2022
Nouvelle fonction	Amazon Kendra La fonction de recherche tabulaire permet de rechercher et d'extraire des réponses à partir de tableaux intégrés dans des documents HTML.	27 novembre 2022
Nouvelle fonction	Amazon Kendra prend en charge la recherche sémantique pour un ensemble sélectionné de langues.	27 novembre 2022
Nouvelle fonction	Amazon Kendra fournit désormais un connecteur de source de données pour Dropbox. Pour plus d'informations, consultez Dropbox.	27 septembre 2022

Nouvelle fonction	Amazon Kendra fournit désormais un connecteur de source de données pour Zendesk. Pour plus d'informations, consultez Zendesk.	17 août 2022
Nouvelle fonction	Le contrôle d'accès au niveau des documents peut désormais être reconfiguré une fois que vous avez indexé vos documents. Pour plus d'informations, consultez la section Configuration du contrôle d'accès.	14 juillet 2022
Nouvelle fonction	Amazon Kendra fournit désormais un connecteur de source de données pour Alfresco. Pour plus d'informations, consultez Alfresco.	30 juin 2022
Nouvelle fonction	Amazon Kendra fournit désormais un connecteur de source de données pour GitHub. Pour de plus amples informations, veuillez consulter GitHub.	2 juin 2022
Nouvelle fonction	Amazon Kendra fournit désormais un connecteur de source de données pour Jira. Pour plus d'informations, consultez Jira.	12 mai 2022

Nouvelle fonction	Les facettes imbriquées dans une facette peuvent être affichées dans les résultats de recherche. Pour plus d'informations, consultez Facets.	5 mai 2022
Nouvelle fonction	Amazon Kendra fournit désormais un connecteur de source de données pour Quip. Pour plus d'informations, consultez Quip.	19 avril 2022
Nouvelle fonction	Amazon Kendra fournit désormais un connecteur de source de données pour Box. Pour plus d'informations, consultez l'encadré.	6 avril 2022
Nouvelle fonction	Amazon Kendra fournit désormais un connecteur de source de données pour Slack. Pour plus d'informations, consultez Slack.	14 mars 2022
Nouvelle fonction	Amazon Kendra fournit désormais un connecteur de source de données pour Amazon FSx. Pour de plus amples informations, veuillez consulter Amazon FSx.	8 février 2022
AWS mises à jour des politiques gérées - Nouvelles politiques	Amazon Kendra a ajouté de nouvelles politiques AWS gérées. Pour plus d'informations, consultez la section Politiques AWS gérées pour Amazon Kendra.	3 janvier 2022

Nouvelle fonction	Amazon Kendra une application de recherche peut être déployée en quelques clics sans avoir besoin de code frontal. Pour plus d'informations, voir Déploiement d'une application de recherche sans code.	1er décembre 2021
Nouvelle fonction	Les métadonnées et le contenu des documents peuvent être enrichis au cours du processus d'ingestion des documents. Pour plus d'informations, veuillez consulter la rubrique Personnalisation des métadonnées de documents pendant le processus d'intégration.	1er décembre 2021
Nouvelle fonction	Amazon Kendra propose des analyses de recherche pour obtenir des informations utiles sur votre application de recherche. Pour plus d'informations, consultez la section Obtenir des informations grâce à l'analyse des recherches.	1er décembre 2021
Expansion de région	Amazon Kendra est désormais disponible en AWS GovCloud (US-West) (us-gov-west-1).	13 octobre 2021

Nouvelle fonction	Amazon Kendra peut désormais indexer des documents dans plusieurs langues et filtrer les résultats de recherche par langue. Consultez les sections Ajout de documents dans des langues autres que l'anglais et Recherche dans les langues.	7 octobre 2021
Nouvelle fonction	Amazon Kendra s'intègre désormais au répertoire Identity Center pour récupérer les niveaux d'accès des groupes et des utilisateurs afin de filtrer le contexte utilisateur. Voir Configuration des groupes d'utilisateurs pour IAM Identity Center.	6 octobre 2021
Nouveau tutoriel	Amazon Kendra propose désormais un didacticiel qui explique comment créer une solution de recherche enrichie de métadonnées. Consultez la section Création d'une solution de recherche intelligente.	13 août 2021
Nouvelle fonction	Amazon Kendra fournit désormais un connecteur de source de données pour WorkDocs. Pour de plus amples informations, veuillez consulter WorkDocs.	20 juillet 2021

Nouvelle fonction	Amazon Kendra fournit désormais un robot d'exploration Web pour explorer et indexer les pages Web. Pour plus d'informations, consultez Web crawler.	17 juin 2021
Expansion de région	Amazon Kendra est maintenant disponible au Canada (centre) (ca-central-1).	16 juin 2021
Expansion de région	Amazon Kendra est désormais disponible dans l'est des États-Unis (Ohio) (us-east-2).	7 juin 2021
Nouvelle fonction	Amazon Kendra prend désormais en charge les suggestions de requêtes, dans lesquelles les utilisateurs se voient proposer des requêtes populaires pertinentes pour leur recherche. Pour plus d'informations, consultez la section Suggestion de requêtes de recherche populaires.	27 mai 2021
AWS mises à jour des politiques gérées - Nouvelles politiques	Amazon Kendra a ajouté de nouvelles politiques AWS gérées. Pour plus d'informations, consultez la section Politiques AWS gérées pour Amazon Kendra.	27 mai 2021
Expansion de région	Amazon Kendra est désormais disponible en Asie-Pacifique (Singapour) (ap-southeast-1).	5 mai 2021

Nouvelle fonction	Amazon Kendra permet désormais de régler la pertinence de la recherche dans la requête en remplaçant les configurations de réglage définies au niveau de l'index. Pour plus d'informations, voir Réglage de la pertinence de la recherche et Réglage des réponses.	20 avril 2021
Nouvelle fonction	Amazon Kendra prend désormais en charge l'authentification OAuth 2.0 et l'utilisation de ServiceNow requêtes pour sélectionner les documents à indexer. Pour de plus amples informations, veuillez consulter ServiceNow.	1 avril 2021
Nouvelle fonction	Amazon Kendra prend désormais en charge l'apprentissage progressif pour les documents de FAQ. Pour plus d'informations, consultez la section Soumission de commentaires pour un apprentissage progressif.	17 février 2021
Nouvelle fonction	Amazon Kendra supporte désormais les synonymes d'index. Pour plus d'informations, consultez la section Ajout de synonymes à un index.	10 décembre 2020

Nouvelle fonction	Amazon Kendra fournit désormais un connecteur de base de données pour Google Workspace Drive. Pour de plus amples informations, consultez Utilisation d'une source de données Google Workspace Drive.	8 décembre 2020
Nouvelle fonction	Amazon Kendra fournit désormais une JavaScript bibliothèque qui vous permet de fournir plus facilement des commentaires sur les requêtes à Amazon Kendra. Pour plus d'informations, consultez la section Soumission de commentaires.	8 décembre 2020
Nouvelle fonction	Amazon Kendra prend désormais en charge le contrôle d'accès utilisateur basé sur des jetons. Pour plus d'informations, consultez la section Contrôle de l'accès aux documents d'un index.	5 novembre 2020
Nouvelle fonction	Le connecteur de source de données Amazon Kendra Confluence fonctionne désormais avec le cloud Confluence. Pour plus d'informations, veuillez consulter la rubrique Utilisation d'une source de données Confluence.	5 novembre 2020

Expansion de région	Amazon Kendra est désormais disponible en Asie-Pacifique (Sydney) (ap-southeast-2).	2 novembre 2020
Nouvelle fonction	Amazon Kendra fournit désormais un connecteur de source de données pour le serveur Confluence. Pour plus d'informations, veuillez consulter la rubrique Utilisation d'une source de données Confluence.	26 octobre 2020
Nouvelle fonction	Amazon Kendra fournit désormais une source de données que vous pouvez utiliser pour générer des statistiques pour vos connecteurs personnalisés. Pour plus d'informations, consultez la section Utilisation d'une source de données personnalisée.	21 octobre 2020
Nouvelle fonction	Amazon Kendra prend désormais en charge les attributs personnalisés pour les questions fréquemment posées. Pour plus d'informations, consultez la section Ajouter des questions et réponses.	17 septembre 2020

Nouvelle fonction	Amazon Kendra renvoie désormais les scores de confiance pour les résultats des requêtes. Pour de plus amples informations, veuillez consulter QueryResultItem.	15 septembre 2020
Nouvelle fonction	AWS CloudFormation prend désormais en charge Amazon Kendra. Pour plus d'informations, voir la référence du type de Amazon Kendra ressource - AWS CloudFormation.	10 septembre 2020
Nouvelle fonction	Amazon Kendra ajoute le support pour AWS PrivateLink. Pour en savoir plus, consultez Amazon Kendra et points de terminaison VPC d’interface (AWS PrivateLink).	7 juillet 2020
Nouveau guide	Il s’agit de la première version du Guide du développeur Amazon Kendra .	11 mai 2020

Les traductions sont fournies par des outils de traduction automatique. En cas de conflit entre le contenu d'une traduction et celui de la version originale en anglais, la version anglaise prévaudra.