



API Reference

Amazon Bedrock AgentCore Data Plane



API Version 2024-02-28

Copyright © 2026 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Amazon Bedrock AgentCore Data Plane: API Reference

Copyright © 2026 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Amazon's trademarks and trade dress may not be used in connection with any product or service that is not Amazon's, in any manner that is likely to cause confusion among customers, or in any manner that disparages or discredits Amazon. All other trademarks not owned by Amazon are the property of their respective owners, who may or may not be affiliated with, connected to, or sponsored by Amazon.

Table of Contents

Welcome	1
Actions	2
BatchCreateMemoryRecords	5
Request Syntax	5
URI Request Parameters	5
Request Body	6
Response Syntax	6
Response Elements	7
Errors	7
See Also	8
BatchDeleteMemoryRecords	10
Request Syntax	10
URI Request Parameters	10
Request Body	10
Response Syntax	11
Response Elements	11
Errors	12
See Also	13
BatchUpdateMemoryRecords	14
Request Syntax	14
URI Request Parameters	14
Request Body	15
Response Syntax	15
Response Elements	16
Errors	16
See Also	17
CompleteResourceTokenAuth	18
Request Syntax	18
URI Request Parameters	18
Request Body	18
Response Syntax	19
Response Elements	19
Errors	19
See Also	20

CreateABTest	22
Request Syntax	22
URI Request Parameters	23
Request Body	23
Response Syntax	25
Response Elements	26
Errors	27
See Also	28
CreateEvent	30
Request Syntax	30
URI Request Parameters	30
Request Body	31
Response Syntax	33
Response Elements	33
Errors	34
See Also	35
CreatePaymentInstrument	36
Request Syntax	36
URI Request Parameters	36
Request Body	36
Response Syntax	38
Response Elements	38
Errors	38
See Also	40
CreatePaymentSession	41
Request Syntax	41
URI Request Parameters	41
Request Body	41
Response Syntax	42
Response Elements	43
Errors	43
See Also	45
DeleteABTest	46
Request Syntax	46
URI Request Parameters	46
Request Body	46

Response Syntax	46
Response Elements	46
Errors	47
See Also	48
DeleteBatchEvaluation	50
Request Syntax	50
URI Request Parameters	50
Request Body	50
Response Syntax	50
Response Elements	50
Errors	51
See Also	52
DeleteEvent	54
Request Syntax	54
URI Request Parameters	54
Request Body	55
Response Syntax	55
Response Elements	55
Errors	55
See Also	57
DeleteMemoryRecord	58
Request Syntax	58
URI Request Parameters	58
Request Body	58
Response Syntax	59
Response Elements	59
Errors	59
See Also	60
DeletePaymentInstrument	62
Request Syntax	62
URI Request Parameters	62
Request Body	62
Response Syntax	63
Response Elements	63
Errors	64
See Also	65

DeletePaymentSession	66
Request Syntax	66
URI Request Parameters	66
Request Body	66
Response Syntax	67
Response Elements	67
Errors	67
See Also	68
DeleteRecommendation	70
Request Syntax	70
URI Request Parameters	70
Request Body	70
Response Syntax	70
Response Elements	70
Errors	71
See Also	72
Evaluate	74
Request Syntax	74
URI Request Parameters	74
Request Body	75
Response Syntax	75
Response Elements	76
Errors	76
See Also	78
GetABTest	80
Request Syntax	80
URI Request Parameters	80
Request Body	80
Response Syntax	80
Response Elements	82
Errors	85
See Also	86
GetAgentCard	88
Request Syntax	88
URI Request Parameters	88
Request Body	88

Response Syntax	88
Response Elements	89
Errors	89
See Also	90
GetBatchEvaluation	92
Request Syntax	92
URI Request Parameters	92
Request Body	92
Response Syntax	92
Response Elements	95
Errors	98
See Also	99
GetBrowserSession	101
Request Syntax	101
URI Request Parameters	101
Request Body	101
Response Syntax	102
Response Elements	103
Errors	105
See Also	106
GetCodeInterpreterSession	108
Request Syntax	108
URI Request Parameters	108
Request Body	109
Response Syntax	109
Response Elements	109
Errors	110
See Also	111
GetEvent	113
Request Syntax	113
URI Request Parameters	113
Request Body	114
Response Syntax	114
Response Elements	114
Errors	115
See Also	116

GetMemoryRecord	117
Request Syntax	117
URI Request Parameters	117
Request Body	117
Response Syntax	117
Response Elements	118
Errors	118
See Also	119
GetPaymentInstrument	121
Request Syntax	121
URI Request Parameters	121
Request Body	121
Response Syntax	122
Response Elements	123
Errors	123
See Also	124
GetPaymentInstrumentBalance	125
Request Syntax	125
URI Request Parameters	125
Request Body	125
Response Syntax	127
Response Elements	127
Errors	128
See Also	129
GetPaymentSession	130
Request Syntax	130
URI Request Parameters	130
Request Body	130
Response Syntax	131
Response Elements	132
Errors	132
See Also	133
GetRecommendation	134
Request Syntax	134
URI Request Parameters	134
Request Body	134

Response Syntax	134
Response Elements	135
Errors	137
See Also	138
GetResourceApiKey	139
Request Syntax	139
URI Request Parameters	139
Request Body	139
Response Syntax	140
Response Elements	140
Errors	140
See Also	141
GetResourceOAuth2Token	143
Request Syntax	143
URI Request Parameters	143
Request Body	143
Response Syntax	146
Response Elements	146
Errors	147
See Also	149
GetResourcePaymentToken	150
Request Syntax	150
URI Request Parameters	150
Request Body	150
Response Syntax	151
Response Elements	151
Errors	151
See Also	153
GetWorkloadAccessToken	154
Request Syntax	154
URI Request Parameters	154
Request Body	154
Response Syntax	154
Response Elements	155
Errors	155
See Also	156

GetWorkloadAccessTokenForJWT	158
Request Syntax	158
URI Request Parameters	158
Request Body	158
Response Syntax	159
Response Elements	159
Errors	159
See Also	160
GetWorkloadAccessTokenForUserId	162
Request Syntax	162
URI Request Parameters	162
Request Body	162
Response Syntax	163
Response Elements	163
Errors	163
See Also	164
InvokeAgentRuntime	166
Request Syntax	166
URI Request Parameters	167
Request Body	169
Response Syntax	169
Response Elements	169
Errors	171
See Also	172
InvokeAgentRuntimeCommand	174
Request Syntax	174
URI Request Parameters	174
Request Body	176
Response Syntax	176
Response Elements	177
Errors	180
See Also	181
InvokeBrowser	183
Request Syntax	183
URI Request Parameters	183
Request Body	184

Response Syntax	184
Response Elements	184
Errors	185
See Also	186
InvokeCodeInterpreter	187
Request Syntax	187
URI Request Parameters	188
Request Body	188
Response Syntax	189
Response Elements	190
Errors	192
See Also	194
InvokeHarness	195
Request Syntax	195
URI Request Parameters	196
Request Body	196
Response Syntax	198
Response Elements	199
Errors	201
See Also	202
ListABTests	204
Request Syntax	204
URI Request Parameters	204
Request Body	204
Response Syntax	204
Response Elements	205
Errors	205
See Also	206
ListActors	208
Request Syntax	208
URI Request Parameters	208
Request Body	208
Response Syntax	209
Response Elements	209
Errors	210
See Also	211

ListBatchEvaluations	212
Request Syntax	212
URI Request Parameters	212
Request Body	212
Response Syntax	212
Response Elements	213
Errors	214
See Also	215
ListBrowserSessions	216
Request Syntax	216
URI Request Parameters	216
Request Body	217
Response Syntax	217
Response Elements	218
Errors	218
See Also	219
ListCodeInterpreterSessions	221
Request Syntax	221
URI Request Parameters	221
Request Body	222
Response Syntax	222
Response Elements	223
Errors	223
See Also	224
ListEvents	226
Request Syntax	226
URI Request Parameters	226
Request Body	227
Response Syntax	228
Response Elements	229
Errors	229
See Also	230
ListMemoryExtractionJobs	232
Request Syntax	232
URI Request Parameters	232
Request Body	232

Response Syntax	233
Response Elements	234
Errors	234
See Also	235
ListMemoryRecords	237
Request Syntax	237
URI Request Parameters	237
Request Body	238
Response Syntax	239
Response Elements	240
Errors	240
See Also	241
ListPaymentInstruments	243
Request Syntax	243
URI Request Parameters	243
Request Body	243
Response Syntax	244
Response Elements	245
Errors	245
See Also	246
ListPaymentSessions	248
Request Syntax	248
URI Request Parameters	248
Request Body	248
Response Syntax	249
Response Elements	250
Errors	250
See Also	251
ListRecommendations	252
Request Syntax	252
URI Request Parameters	252
Request Body	252
Response Syntax	253
Response Elements	253
Errors	254
See Also	254

ListSessions	256
Request Syntax	256
URI Request Parameters	256
Request Body	257
Response Syntax	257
Response Elements	258
Errors	258
See Also	259
ProcessPayment	261
Request Syntax	261
URI Request Parameters	261
Request Body	261
Response Syntax	263
Response Elements	263
Errors	265
See Also	266
RetrieveMemoryRecords	268
Request Syntax	268
URI Request Parameters	268
Request Body	269
Response Syntax	270
Response Elements	271
Errors	271
See Also	272
SaveBrowserSessionProfile	274
Request Syntax	274
URI Request Parameters	274
Request Body	275
Response Syntax	276
Response Elements	276
Errors	277
See Also	278
SearchRegistryRecords	280
Request Syntax	280
URI Request Parameters	280
Request Body	280

Response Syntax	281
Response Elements	282
Errors	283
See Also	284
StartBatchEvaluation	285
Request Syntax	285
URI Request Parameters	285
Request Body	285
Response Syntax	288
Response Elements	289
Errors	291
See Also	292
StartBrowserSession	294
Request Syntax	294
URI Request Parameters	295
Request Body	296
Response Syntax	298
Response Elements	298
Errors	299
See Also	300
StartCodeInterpreterSession	302
Request Syntax	302
URI Request Parameters	302
Request Body	303
Response Syntax	304
Response Elements	304
Errors	305
See Also	306
StartMemoryExtractionJob	308
Request Syntax	308
URI Request Parameters	308
Request Body	308
Response Syntax	309
Response Elements	309
Errors	309
See Also	311

StartRecommendation	312
Request Syntax	312
URI Request Parameters	312
Request Body	312
Response Syntax	314
Response Elements	315
Errors	316
See Also	317
StopBatchEvaluation	319
Request Syntax	319
URI Request Parameters	319
Request Body	319
Response Syntax	319
Response Elements	319
Errors	320
See Also	322
StopBrowserSession	323
Request Syntax	323
URI Request Parameters	323
Request Body	324
Response Syntax	324
Response Elements	325
Errors	325
See Also	326
StopCodeInterpreterSession	328
Request Syntax	328
URI Request Parameters	328
Request Body	329
Response Syntax	329
Response Elements	330
Errors	330
See Also	331
StopRuntimeSession	333
Request Syntax	333
URI Request Parameters	333
Request Body	333

Response Syntax	334
Response Elements	334
Errors	334
See Also	336
UpdateABTest	338
Request Syntax	338
URI Request Parameters	338
Request Body	339
Response Syntax	341
Response Elements	341
Errors	342
See Also	343
UpdateBrowserStream	345
Request Syntax	345
URI Request Parameters	345
Request Body	345
Response Syntax	346
Response Elements	346
Errors	347
See Also	348
Data Types	350
A2aDescriptor	359
Contents	359
See Also	359
ABTestEvaluationConfig	360
Contents	360
See Also	360
ABTestResults	362
Contents	362
See Also	362
ABTestSummary	363
Contents	363
See Also	365
ActorSummary	366
Contents	366
See Also	366

AffectedSession	367
Contents	367
See Also	368
AgentCardDefinition	369
Contents	369
See Also	369
AgentSkillsDescriptor	370
Contents	370
See Also	370
AgentTracesConfig	371
Contents	371
See Also	371
Amount	373
Contents	373
See Also	373
AutomationStream	374
Contents	374
See Also	374
AutomationStreamUpdate	375
Contents	375
See Also	375
AvailableLimits	376
Contents	376
See Also	376
BasicAuth	377
Contents	377
See Also	377
BatchEvaluationSummary	378
Contents	378
See Also	380
BatchEvaluationTraceConfig	381
Contents	381
See Also	381
Branch	382
Contents	382
See Also	382

BranchFilter	383
Contents	383
See Also	383
BrowserAction	384
Contents	384
See Also	385
BrowserActionResult	386
Contents	386
See Also	387
BrowserEnterprisePolicy	388
Contents	388
See Also	388
BrowserExtension	389
Contents	389
See Also	389
BrowserProfileConfiguration	390
Contents	390
See Also	390
BrowserSessionStream	391
Contents	391
See Also	391
BrowserSessionSummary	392
Contents	392
See Also	393
Certificate	394
Contents	394
See Also	394
CertificateLocation	395
Contents	395
See Also	395
CloudWatchFilterConfig	396
Contents	396
See Also	396
CloudWatchLogsFilter	397
Contents	397
See Also	397

CloudWatchLogsRule	399
Contents	399
See Also	399
CloudWatchLogsSource	400
Contents	400
See Also	400
CloudWatchLogsTraceConfig	402
Contents	402
See Also	403
CloudWatchOutputConfig	404
Contents	404
See Also	404
CodeInterpreterResult	405
Contents	405
See Also	405
CodeInterpreterSessionSummary	407
Contents	407
See Also	408
CodeInterpreterStreamOutput	409
Contents	409
See Also	411
CoinbaseCdpTokenRequestInput	412
Contents	412
See Also	413
CoinbaseCdpTokenResponseOutput	414
Contents	414
See Also	414
ConfidenceInterval	415
Contents	415
See Also	415
ConfigurationBundleRef	416
Contents	416
See Also	416
ConfigurationBundleToolEntry	417
Contents	417
See Also	417

Content	418
Contents	418
See Also	418
ContentBlock	419
Contents	419
See Also	420
ContentDeltaEvent	421
Contents	421
See Also	421
ContentStartEvent	422
Contents	422
See Also	422
ContentStopEvent	423
Contents	423
See Also	423
Context	424
Contents	424
See Also	424
ControlStats	425
Contents	425
See Also	425
Conversational	426
Contents	426
See Also	426
CryptoX402PaymentInput	427
Contents	427
See Also	427
CryptoX402PaymentOutput	428
Contents	428
See Also	428
CustomDescriptor	429
Contents	429
See Also	429
DataSourceConfig	430
Contents	430
See Also	430

Descriptors	431
Contents	431
See Also	432
EmbeddedCryptoWallet	433
Contents	433
See Also	434
EvaluationContent	435
Contents	435
See Also	435
EvaluationExpectedTrajectory	436
Contents	436
See Also	436
EvaluationInput	437
Contents	437
See Also	437
EvaluationJobResults	438
Contents	438
See Also	439
EvaluationMetadata	440
Contents	440
See Also	440
EvaluationReferenceInput	441
Contents	441
See Also	442
EvaluationResultContent	443
Contents	443
See Also	446
EvaluationTarget	447
Contents	447
See Also	448
Evaluator	449
Contents	449
See Also	449
EvaluatorMetric	450
Contents	450
See Also	450

EvaluatorStatistics	451
Contents	451
See Also	451
EvaluatorSummary	452
Contents	452
See Also	452
Event	454
Contents	454
See Also	456
EventMetadataFilterExpression	457
Contents	457
See Also	457
ExecutionSummaryAffectedSession	459
Contents	459
See Also	459
ExecutionSummaryCluster	460
Contents	460
See Also	461
ExecutionSummaryClusteringResultContent	462
Contents	462
See Also	462
ExternalProxy	463
Contents	463
See Also	464
ExtractionJob	465
Contents	465
See Also	465
ExtractionJobFilterInput	466
Contents	466
See Also	467
ExtractionJobMessages	468
Contents	468
See Also	468
ExtractionJobMetadata	469
Contents	469
See Also	470

FailureAnalysisResultContent	471
Contents	471
See Also	471
FailureCategoryCluster	472
Contents	472
See Also	473
FailureSpanDetail	474
Contents	474
See Also	474
FailureSubCategoryCluster	475
Contents	475
See Also	476
FilterInput	477
Contents	477
See Also	477
FilterValue	478
Contents	478
See Also	478
GatewayFilter	480
Contents	480
See Also	480
GroundTruthSource	481
Contents	481
See Also	481
GroundTruthTurn	482
Contents	482
See Also	482
GroundTruthTurnInput	483
Contents	483
See Also	483
HarnessAgentCoreBrowserConfig	484
Contents	484
See Also	484
HarnessAgentCoreCodeInterpreterConfig	485
Contents	485
See Also	485

HarnessAgentCoreGatewayConfig	486
Contents	486
See Also	486
HarnessBedrockModelConfig	487
Contents	487
See Also	488
HarnessContentBlock	489
Contents	489
See Also	490
HarnessContentBlockDelta	491
Contents	491
See Also	492
HarnessContentBlockDeltaEvent	493
Contents	493
See Also	493
HarnessContentBlockStart	494
Contents	494
See Also	494
HarnessContentBlockStartEvent	495
Contents	495
See Also	495
HarnessContentBlockStopEvent	496
Contents	496
See Also	496
HarnessGatewayOutboundAuth	497
Contents	497
See Also	497
HarnessGeminiModelConfig	499
Contents	499
See Also	500
HarnessInlineFunctionConfig	501
Contents	501
See Also	501
HarnessLiteLlmModelConfig	502
Contents	502
See Also	503

HarnessMessage	504
Contents	504
See Also	504
HarnessMessageStartEvent	505
Contents	505
See Also	505
HarnessMessageStopEvent	506
Contents	506
See Also	506
HarnessMetadataEvent	507
Contents	507
See Also	507
HarnessModelConfiguration	508
Contents	508
See Also	509
HarnessOpenAiModelConfig	510
Contents	510
See Also	511
HarnessReasoningContentBlock	512
Contents	512
See Also	512
HarnessReasoningContentBlockDelta	513
Contents	513
See Also	513
HarnessReasoningTextBlock	515
Contents	515
See Also	515
HarnessRemoteMcpConfig	516
Contents	516
See Also	516
HarnessSkill	517
Contents	517
See Also	518
HarnessSkillAwsSkillsSource	519
Contents	519
See Also	519

HarnessSkillGitAuth	520
Contents	520
See Also	520
HarnessSkillGitSource	521
Contents	521
See Also	521
HarnessSkillsS3Source	523
Contents	523
See Also	523
HarnessStreamMetrics	524
Contents	524
See Also	524
HarnessSystemContentBlock	525
Contents	525
See Also	525
HarnessTokenUsage	526
Contents	526
See Also	527
HarnessTool	528
Contents	528
See Also	528
HarnessToolConfiguration	530
Contents	530
See Also	531
HarnessToolResultBlock	532
Contents	532
See Also	533
HarnessToolResultBlockDelta	534
Contents	534
See Also	534
HarnessToolResultBlockStart	535
Contents	535
See Also	535
HarnessToolResultContentBlock	536
Contents	536
See Also	536

HarnessToolUseBlock	537
Contents	537
See Also	538
HarnessToolUseBlockDelta	539
Contents	539
See Also	539
HarnessToolUseBlockStart	540
Contents	540
See Also	541
InlineGroundTruth	542
Contents	542
See Also	542
InputContentBlock	544
Contents	544
See Also	544
Insight	546
Contents	546
See Also	546
InsightsFailureSignal	547
Contents	547
See Also	548
InvokeAgentRuntimeCommandRequestBody	549
Contents	549
See Also	549
InvokeAgentRuntimeCommandStreamOutput	550
Contents	550
See Also	552
InvokeHarnessStreamOutput	553
Contents	553
See Also	554
KeyPressArguments	556
Contents	556
See Also	556
KeyPressResult	557
Contents	557
See Also	557

KeyShortcutArguments	558
Contents	558
See Also	558
KeyShortcutResult	559
Contents	559
See Also	559
KeyTypeArguments	560
Contents	560
See Also	560
KeyTypeResult	561
Contents	561
See Also	561
LeftExpression	562
Contents	562
See Also	562
LinkedAccount	563
Contents	563
See Also	564
LinkedAccountDeveloperJwt	565
Contents	565
See Also	565
LinkedAccountEmail	566
Contents	566
See Also	566
LinkedAccountOAuth2	567
Contents	567
See Also	568
LinkedAccountSms	569
Contents	569
See Also	569
LiveViewStream	570
Contents	570
See Also	570
McpDescriptor	571
Contents	571
See Also	571

MemoryContent	572
Contents	572
See Also	572
MemoryMetadataFilterExpression	573
Contents	573
See Also	573
MemoryRecord	575
Contents	575
See Also	576
MemoryRecordCreateInput	577
Contents	577
See Also	578
MemoryRecordDeleteInput	579
Contents	579
See Also	579
MemoryRecordLeftExpression	580
Contents	580
See Also	580
MemoryRecordMetadataValue	581
Contents	581
See Also	582
MemoryRecordOutput	583
Contents	583
See Also	584
MemoryRecordRightExpression	585
Contents	585
See Also	585
MemoryRecordSummary	586
Contents	586
See Also	587
MemoryRecordUpdateInput	589
Contents	589
See Also	590
MessageMetadata	591
Contents	591
See Also	591

MetadataValue	592
Contents	592
See Also	592
MouseClickedArguments	593
Contents	593
See Also	593
MouseClickedResult	595
Contents	595
See Also	595
MouseDownArguments	596
Contents	596
See Also	597
MouseDownResult	598
Contents	598
See Also	598
MouseMoveArguments	599
Contents	599
See Also	599
MouseMoveResult	600
Contents	600
See Also	600
MouseScrollArguments	601
Contents	601
See Also	601
MouseScrollResult	603
Contents	603
See Also	603
OAuth2Authentication	604
Contents	604
See Also	605
OAuthCredentialProvider	606
Contents	606
See Also	607
OnlineEvaluationConfigSource	608
Contents	608
See Also	608

OutputConfig	609
Contents	609
See Also	609
PayloadType	610
Contents	610
See Also	610
PaymentInput	611
Contents	611
See Also	611
PaymentInstrument	612
Contents	612
See Also	614
PaymentInstrumentDetails	615
Contents	615
See Also	615
PaymentInstrumentSummary	616
Contents	616
See Also	618
PaymentOutput	619
Contents	619
See Also	619
PaymentSession	620
Contents	620
See Also	621
PaymentSessionSummary	623
Contents	623
See Also	624
PaymentTokenRequestInput	625
Contents	625
See Also	625
PaymentTokenResponseOutput	626
Contents	626
See Also	626
PerVariantOnlineEvaluationConfig	627
Contents	627
See Also	627

Proxy	628
Contents	628
See Also	628
ProxyBypass	629
Contents	629
See Also	629
ProxyConfiguration	630
Contents	630
See Also	630
ProxyCredentials	631
Contents	631
See Also	631
RecommendationConfig	632
Contents	632
See Also	632
RecommendationEvaluationConfig	633
Contents	633
See Also	633
RecommendationEvaluatorReference	634
Contents	634
See Also	634
RecommendationResult	635
Contents	635
See Also	635
RecommendationResultConfigurationBundle	636
Contents	636
See Also	636
RecommendationSummary	637
Contents	637
See Also	638
RegistryRecordSummary	640
Contents	640
See Also	642
ResourceContent	643
Contents	643
See Also	644

ResourceLocation	645
Contents	645
See Also	645
ResponseChunk	646
Contents	646
See Also	646
RightExpression	648
Contents	648
See Also	648
RootCauseCluster	649
Contents	649
See Also	650
S3Location	651
Contents	651
See Also	651
ScreenshotArguments	653
Contents	653
See Also	653
ScreenshotResult	654
Contents	654
See Also	654
SearchCriteria	655
Contents	655
See Also	656
SecretsManagerLocation	657
Contents	657
See Also	657
ServerDefinition	658
Contents	658
See Also	658
SessionFilter	659
Contents	659
See Also	659
SessionFilterConfig	660
Contents	660
See Also	660

SessionLimits	661
Contents	661
See Also	661
SessionMetadataShape	662
Contents	662
See Also	662
SessionSummary	664
Contents	664
See Also	664
SkillDefinition	666
Contents	666
See Also	666
SkillMdDefinition	667
Contents	667
See Also	667
SpanContext	668
Contents	668
See Also	668
StreamUpdate	670
Contents	670
See Also	670
StripePrivyTokenRequestInput	671
Contents	671
See Also	672
StripePrivyTokenResponseOutput	673
Contents	673
See Also	674
SystemPromptConfig	675
Contents	675
See Also	675
SystemPromptConfigurationBundle	676
Contents	676
See Also	676
SystemPromptRecommendationConfig	678
Contents	678
See Also	678

SystemPromptRecommendationResult	680
Contents	680
See Also	681
TargetRef	682
Contents	682
See Also	682
TokenBalance	683
Contents	683
See Also	684
TokenUsage	685
Contents	685
See Also	685
ToolArguments	687
Contents	687
See Also	689
ToolDescriptionConfig	690
Contents	690
See Also	690
ToolDescriptionConfigurationBundle	691
Contents	691
See Also	691
ToolDescriptionInput	693
Contents	693
See Also	693
ToolDescriptionOutput	694
Contents	694
See Also	694
ToolDescriptionRecommendationConfig	696
Contents	696
See Also	696
ToolDescriptionRecommendationResult	697
Contents	697
See Also	698
ToolDescriptionSource	699
Contents	699
See Also	699

ToolDescriptionTextInput	700
Contents	700
See Also	700
ToolResultStructuredContent	701
Contents	701
See Also	702
ToolsDefinition	703
Contents	703
See Also	703
UserIdentifier	704
Contents	704
See Also	704
UserIntentAffectedSession	706
Contents	706
See Also	706
UserIntentCluster	707
Contents	707
See Also	708
UserIntentClusteringResultContent	709
Contents	709
See Also	709
ValidationExceptionField	710
Contents	710
See Also	710
Variant	711
Contents	711
See Also	711
VariantConfiguration	713
Contents	713
See Also	713
VariantResult	714
Contents	714
See Also	715
ViewPort	716
Contents	716
See Also	716

Common Parameters	717
Common Error Types	720

Welcome

Welcome to the Amazon Bedrock AgentCore Data Plane API reference. Data Plane actions process and handle data or workloads within AWS services.

This document was last published on June 26, 2026.

Actions

The following actions are supported:

- [BatchCreateMemoryRecords](#)
- [BatchDeleteMemoryRecords](#)
- [BatchUpdateMemoryRecords](#)
- [CompleteResourceTokenAuth](#)
- [CreateABTest](#)
- [CreateEvent](#)
- [CreatePaymentInstrument](#)
- [CreatePaymentSession](#)
- [DeleteABTest](#)
- [DeleteBatchEvaluation](#)
- [DeleteEvent](#)
- [DeleteMemoryRecord](#)
- [DeletePaymentInstrument](#)
- [DeletePaymentSession](#)
- [DeleteRecommendation](#)
- [Evaluate](#)
- [GetABTest](#)
- [GetAgentCard](#)
- [GetBatchEvaluation](#)
- [GetBrowserSession](#)
- [GetCodeInterpreterSession](#)
- [GetEvent](#)
- [GetMemoryRecord](#)
- [GetPaymentInstrument](#)
- [GetPaymentInstrumentBalance](#)
- [GetPaymentSession](#)
- [GetRecommendation](#)

- [GetResourceApiKey](#)
- [GetResourceOauth2Token](#)
- [GetResourcePaymentToken](#)
- [GetWorkloadAccessToken](#)
- [GetWorkloadAccessTokenForJWT](#)
- [GetWorkloadAccessTokenForUserId](#)
- [InvokeAgentRuntime](#)
- [InvokeAgentRuntimeCommand](#)
- [InvokeBrowser](#)
- [InvokeCodeInterpreter](#)
- [InvokeHarness](#)
- [ListABTests](#)
- [ListActors](#)
- [ListBatchEvaluations](#)
- [ListBrowserSessions](#)
- [ListCodeInterpreterSessions](#)
- [ListEvents](#)
- [ListMemoryExtractionJobs](#)
- [ListMemoryRecords](#)
- [ListPaymentInstruments](#)
- [ListPaymentSessions](#)
- [ListRecommendations](#)
- [ListSessions](#)
- [ProcessPayment](#)
- [RetrieveMemoryRecords](#)
- [SaveBrowserSessionProfile](#)
- [SearchRegistryRecords](#)
- [StartBatchEvaluation](#)
- [StartBrowserSession](#)
- [StartCodeInterpreterSession](#)

- [StartMemoryExtractionJob](#)
- [StartRecommendation](#)
- [StopBatchEvaluation](#)
- [StopBrowserSession](#)
- [StopCodeInterpreterSession](#)
- [StopRuntimeSession](#)
- [UpdateABTest](#)
- [UpdateBrowserStream](#)

BatchCreateMemoryRecords

Creates multiple memory records in a single batch operation for the specified memory with custom content.

Request Syntax

```
POST /memories/memoryId/memoryRecords/batchCreate HTTP/1.1
```

```
Content-type: application/json
```

```
{
  "clientToken": "string",
  "records": [
    {
      "content": { ... },
      "memoryStrategyId": "string",
      "metadata": {
        "string" : { ... }
      },
      "namespaces": [ "string" ],
      "requestIdentifier": "string",
      "timestamp": number
    }
  ]
}
```

URI Request Parameters

The request uses the following URI parameters.

memoryId

The unique ID of the memory resource where records will be created.

Length Constraints: Minimum length of 12.

Pattern: (arn:(aws|aws-cn|aws-us-gov):bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:memory/)?[a-zA-Z][a-zA-Z0-9-]{0,99}-[a-zA-Z0-9]{10}

Required: Yes

Request Body

The request accepts the following data in JSON format.

clientToken

A unique, case-sensitive identifier to ensure idempotent processing of the batch request.

Type: String

Required: No

records

A list of memory record creation inputs to be processed in the batch operation.

Type: Array of [MemoryRecordCreateInput](#) objects

Array Members: Minimum number of 0 items. Maximum number of 100 items.

Required: Yes

Response Syntax

```
HTTP/1.1 201
Content-type: application/json

{
  "failedRecords": [
    {
      "errorCode": number,
      "errorMessage": "string",
      "memoryRecordId": "string",
      "requestIdentifier": "string",
      "status": "string"
    }
  ],
  "successfulRecords": [
    {
      "errorCode": number,
      "errorMessage": "string",
      "memoryRecordId": "string",
```

```
    "requestIdentifier": "string",  
    "status": "string"  
  }  
]  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 201 response.

The following data is returned in JSON format by the service.

failedRecords

A list of memory records that failed to be created, including error details for each failure.

Type: Array of [MemoryRecordOutput](#) objects

successfulRecords

A list of memory records that were successfully created during the batch operation.

Type: Array of [MemoryRecordOutput](#) objects

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ServiceException

The service encountered an internal error. Try your request again later.

HTTP Status Code: 500

ServiceQuotaExceededException

The exception that occurs when the request would cause a service quota to be exceeded. Review your service quotas and either reduce your request rate or request a quota increase.

HTTP Status Code: 402

ThrottledException

The request was denied due to request throttling. Reduce the frequency of requests and try again.

HTTP Status Code: 429

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)

- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

BatchDeleteMemoryRecords

Deletes multiple memory records in a single batch operation from the specified memory.

Request Syntax

```
POST /memories/memoryId/memoryRecords/batchDelete HTTP/1.1
Content-type: application/json
```

```
{
  "records": [
    {
      "memoryRecordId": "string"
    }
  ]
}
```

URI Request Parameters

The request uses the following URI parameters.

[memoryId](#)

The unique ID of the memory resource where records will be deleted.

Length Constraints: Minimum length of 12.

Pattern: (arn:(aws|aws-cn|aws-us-gov):bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:memory/)?[a-zA-Z][a-zA-Z0-9-_{0,99}]-[a-zA-Z0-9]{10}

Required: Yes

Request Body

The request accepts the following data in JSON format.

[records](#)

A list of memory record deletion inputs to be processed in the batch operation.

Type: Array of [MemoryRecordDeleteInput](#) objects

Array Members: Minimum number of 0 items. Maximum number of 100 items.

Required: Yes

Response Syntax

```
HTTP/1.1 200
Content-type: application/json

{
  "failedRecords": [
    {
      "errorCode": number,
      "errorMessage": "string",
      "memoryRecordId": "string",
      "requestIdentifier": "string",
      "status": "string"
    }
  ],
  "successfulRecords": [
    {
      "errorCode": number,
      "errorMessage": "string",
      "memoryRecordId": "string",
      "requestIdentifier": "string",
      "status": "string"
    }
  ]
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

failedRecords

A list of memory records that failed to be deleted, including error details for each failure.

Type: Array of [MemoryRecordOutput](#) objects

[successfulRecords](#)

A list of memory records that were successfully deleted during the batch operation.

Type: Array of [MemoryRecordOutput](#) objects

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ServiceException

The service encountered an internal error. Try your request again later.

HTTP Status Code: 500

ServiceQuotaExceededException

The exception that occurs when the request would cause a service quota to be exceeded. Review your service quotas and either reduce your request rate or request a quota increase.

HTTP Status Code: 402

ThrottledException

The request was denied due to request throttling. Reduce the frequency of requests and try again.

HTTP Status Code: 429

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

BatchUpdateMemoryRecords

Updates multiple memory records with custom content in a single batch operation within the specified memory.

Request Syntax

```
POST /memories/memoryId/memoryRecords/batchUpdate HTTP/1.1
```

```
Content-type: application/json
```

```
{
  "records": [
    {
      "content": { ... },
      "memoryRecordId": "string",
      "memoryStrategyId": "string",
      "metadata": {
        "string" : { ... }
      },
      "namespaces": [ "string" ],
      "timestamp": number
    }
  ]
}
```

URI Request Parameters

The request uses the following URI parameters.

memoryId

The unique ID of the memory resource where records will be updated.

Length Constraints: Minimum length of 12.

Pattern: (arn:(aws|aws-cn|aws-us-gov):bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:memory/)?[a-zA-Z][a-zA-Z0-9-_{0,99}]-[a-zA-Z0-9]{10}

Required: Yes

Request Body

The request accepts the following data in JSON format.

records

A list of memory record update inputs to be processed in the batch operation.

Type: Array of [MemoryRecordUpdateInput](#) objects

Array Members: Minimum number of 0 items. Maximum number of 100 items.

Required: Yes

Response Syntax

```
HTTP/1.1 200
Content-type: application/json

{
  "failedRecords": [
    {
      "errorCode": number,
      "errorMessage": "string",
      "memoryRecordId": "string",
      "requestIdentifier": "string",
      "status": "string"
    }
  ],
  "successfulRecords": [
    {
      "errorCode": number,
      "errorMessage": "string",
      "memoryRecordId": "string",
      "requestIdentifier": "string",
      "status": "string"
    }
  ]
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

failedRecords

A list of memory records that failed to be updated, including error details for each failure.

Type: Array of [MemoryRecordOutput](#) objects

successfulRecords

A list of memory records that were successfully updated during the batch operation.

Type: Array of [MemoryRecordOutput](#) objects

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ServiceException

The service encountered an internal error. Try your request again later.

HTTP Status Code: 500

ServiceQuotaExceededException

The exception that occurs when the request would cause a service quota to be exceeded. Review your service quotas and either reduce your request rate or request a quota increase.

HTTP Status Code: 402

ThrottledException

The request was denied due to request throttling. Reduce the frequency of requests and try again.

HTTP Status Code: 429

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

CompleteResourceTokenAuth

Confirms the user authentication session for obtaining OAuth2.0 tokens for a resource.

Request Syntax

```
POST /identities/CompleteResourceTokenAuth HTTP/1.1
Content-type: application/json

{
  "sessionUri": "string",
  "userIdentifier": { ... }
}
```

URI Request Parameters

The request does not use any URI parameters.

Request Body

The request accepts the following data in JSON format.

sessionUri

Unique identifier for the user's authentication session for retrieving OAuth2 tokens. This ID tracks the authorization flow state across multiple requests and responses during the OAuth2 authentication process.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1024.

Pattern: `urn:ietf:params:oauth:request_uri:[a-zA-Z0-9-._~]+`

Required: Yes

userIdentifier

The OAuth2.0 token or user ID that was used to generate the workload access token used for initiating the user authorization flow to retrieve OAuth2.0 tokens.

Type: [UserIdentifier](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: Yes

Response Syntax

```
HTTP/1.1 200
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

InternalServerError

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

UnauthorizedException

This exception is thrown when the JWT bearer token is invalid or not found for OAuth bearer token based access

HTTP Status Code: 401

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

CreateABTest

Creates an A/B test for comparing agent configurations. A/B tests split traffic between a control variant and a treatment variant through a gateway, then evaluate performance using online evaluation configurations to determine which variant performs better.

Request Syntax

```
POST /ab-tests HTTP/1.1
Content-type: application/json

{
  "clientToken": "string",
  "description": "string",
  "enableOnCreate": boolean,
  "evaluationConfig": { ... },
  "gatewayArn": "string",
  "gatewayFilter": {
    "targetPaths": [ "string" ]
  },
  "name": "string",
  "roleArn": "string",
  "tags": {
    "string" : "string"
  },
  "variants": [
    {
      "name": "string",
      "variantConfiguration": {
        "configurationBundle": {
          "bundleArn": "string",
          "bundleVersion": "string"
        },
        "target": {
          "name": "string"
        }
      },
      "weight": number
    }
  ]
}
```

URI Request Parameters

The request does not use any URI parameters.

Request Body

The request accepts the following data in JSON format.

clientToken

A unique, case-sensitive identifier to ensure that the API request completes no more than one time. If this token matches a previous request, the service ignores the request, but does not return an error.

Type: String

Length Constraints: Minimum length of 33. Maximum length of 256.

Pattern: `[a-zA-Z0-9](-*[a-zA-Z0-9]){0,256}`

Required: No

description

The description of the A/B test.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 200.

Required: No

enableOnCreate

Whether to enable the A/B test immediately upon creation. If true, traffic splitting begins automatically.

Type: Boolean

Required: No

evaluationConfig

The evaluation configuration specifying which online evaluation configurations to use for measuring variant performance.

Type: [ABTestEvaluationConfig](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: Yes

[gatewayArn](#)

The Amazon Resource Name (ARN) of the gateway to use for traffic splitting.

Type: String

Pattern: `arn:aws(|-cn|-us-gov):bedrock-agentcore:[a-z0-9-]{1,20}:[\0-9]{12}:gateway/([\0-9a-z]([-]?){1,48}-[a-z0-9]{10}`

Required: Yes

[gatewayFilter](#)

Optional filter to restrict which gateway target paths are included in the A/B test.

Type: [GatewayFilter](#) object

Required: No

[name](#)

The name of the A/B test. Must be unique within your account.

Type: String

Pattern: `[a-zA-Z][a-zA-Z0-9_]{0,47}`

Required: Yes

[roleArn](#)

The IAM role ARN that grants permissions for the A/B test to access gateway and evaluation resources.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: `arn:aws(-[^\:]+)?:iam::([\0-9]{12})?:role/.+`

Required: Yes

tags

A map of tag keys and values to associate with the A/B test.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 50 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Key Pattern: `[a-zA-Z0-9\s._:/=+@-]*`

Value Length Constraints: Minimum length of 0. Maximum length of 256.

Value Pattern: `[a-zA-Z0-9\s._:/=+@-]*`

Required: No

variants

The list of variants for the A/B test. Must contain exactly two variants: a control (C) and a treatment (T1), each with a configuration bundle or target reference and a traffic weight.

Type: Array of [Variant](#) objects

Array Members: Fixed number of 2 items.

Required: Yes

Response Syntax

```
HTTP/1.1 202
Content-type: application/json

{
  "abTestArn": "string",
  "abTestId": "string",
  "createdAt": number,
  "executionStatus": "string",
  "name": "string",
  "status": "string"
```

```
}
```

Response Elements

If the action is successful, the service sends back an HTTP 202 response.

The following data is returned in JSON format by the service.

abTestArn

The Amazon Resource Name (ARN) of the created A/B test.

Type: String

Pattern: `arn:aws[a-zA-Z-]*:bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:ab-test/[a-zA-Z][a-zA-Z0-9-_{0,99}}-[a-zA-Z0-9]{10}`

abTestId

The unique identifier of the created A/B test.

Type: String

Pattern: `[a-zA-Z][a-zA-Z0-9-_{0,99}}-[a-zA-Z0-9]{10}`

createdAt

The timestamp when the A/B test was created.

Type: Timestamp

executionStatus

The execution status indicating whether the A/B test is currently running.

Type: String

Valid Values: PAUSED | RUNNING | STOPPED | NOT_STARTED

name

The name of the A/B test.

Type: String

Pattern: `[a-zA-Z][a-zA-Z0-9-_{0,47}}`

status

The status of the A/B test.

Type: String

Valid Values: CREATING | ACTIVE | CREATE_FAILED | UPDATING | UPDATE_FAILED | DELETING | DELETE_FAILED | FAILED

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

ConflictException

The exception that occurs when the request conflicts with the current state of the resource. This can happen when trying to modify a resource that is currently being modified by another request, or when trying to create a resource that already exists.

HTTP Status Code: 409

InternalServerError

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ServiceQuotaExceededException

The exception that occurs when the request would cause a service quota to be exceeded. Review your service quotas and either reduce your request rate or request a quota increase.

HTTP Status Code: 402

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

UnauthorizedException

This exception is thrown when the JWT bearer token is invalid or not found for OAuth bearer token based access

HTTP Status Code: 401

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

CreateEvent

Creates an event in an AgentCore Memory resource. Events represent interactions or activities that occur within a session and are associated with specific actors.

To use this operation, you must have the `bedrock-agentcore:CreateEvent` permission.

This operation is subject to request rate limiting.

Request Syntax

```
POST /memories/memoryId/events HTTP/1.1
Content-type: application/json
```

```
{
  "actorId": "string",
  "branch": {
    "name": "string",
    "rootEventId": "string"
  },
  "clientToken": "string",
  "eventTimestamp": number,
  "extractionMode": "string",
  "metadata": {
    "string" : { ... }
  },
  "payload": [
    { ... }
  ],
  "sessionId": "string"
}
```

URI Request Parameters

The request uses the following URI parameters.

memoryId

The identifier of the AgentCore Memory resource in which to create the event.

Length Constraints: Minimum length of 12.

Pattern: `(arn:(aws|aws-cn|aws-us-gov):bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:memory/)?[a-zA-Z][a-zA-Z0-9-_{0,99}]-[a-zA-Z0-9]{10}`

Required: Yes

Request Body

The request accepts the following data in JSON format.

actorId

The identifier of the actor associated with this event. An actor represents an entity that participates in sessions and generates events.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[a-zA-Z0-9][a-zA-Z0-9-_/]*(?::[a-zA-Z0-9-_/]+)*[a-zA-Z0-9-_/]*`

Required: Yes

branch

The branch information for this event. Branches allow for organizing events into different conversation threads or paths.

Type: [Branch](#) object

Required: No

clientToken

A unique, case-sensitive identifier to ensure that the operation completes no more than one time. If this token matches a previous request, AgentCore ignores the request, but does not return an error.

Type: String

Required: No

eventTimestamp

The timestamp when the event occurred. If not specified, the current time is used.

Type: Timestamp

Required: Yes

extractionMode

Controls long-term memory extraction for this event. When set to SKIP, the event is stored in short-term memory but is excluded from long-term memory extraction. If not specified, the event is processed for extraction as usual.

Type: String

Valid Values: SKIP

Required: No

metadata

The key-value metadata to attach to the event.

Type: String to [MetadataValue](#) object map

Map Entries: Minimum number of 0 items. Maximum number of 15 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Key Pattern: `[a-zA-Z0-9\s._:/=+@-]*`

Required: No

payload

The content payload of the event. This can include conversational data or binary content.

Type: Array of [PayloadType](#) objects

Array Members: Minimum number of 0 items. Maximum number of 100 items.

Required: Yes

sessionId

The identifier of the session in which this event occurs. A session represents a sequence of related events.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 100.

Pattern: `[a-zA-Z0-9][a-zA-Z0-9-_*]`

Required: No

Response Syntax

```
HTTP/1.1 201
Content-type: application/json

{
  "event": {
    "actorId": "string",
    "branch": {
      "name": "string",
      "rootEventId": "string"
    },
    "eventId": "string",
    "eventTimestamp": number,
    "memoryId": "string",
    "metadata": {
      "string" : { ... }
    },
    "payload": [
      { ... }
    ],
    "sessionId": "string"
  }
}
```

Response Elements

If the action is successful, the service sends back an HTTP 201 response.

The following data is returned in JSON format by the service.

event

The event that was created.

Type: [Event](#) object

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

InvalidInputException

The input fails to satisfy the constraints specified by AgentCore. Check your input values and try again.

HTTP Status Code: 400

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

RetryableConflictException

The exception that occurs when there is a retryable conflict performing an operation. This is a temporary condition that may resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 409

ServiceException

The service encountered an internal error. Try your request again later.

HTTP Status Code: 500

ServiceQuotaExceededException

The exception that occurs when the request would cause a service quota to be exceeded. Review your service quotas and either reduce your request rate or request a quota increase.

HTTP Status Code: 402

ThrottledException

The request was denied due to request throttling. Reduce the frequency of requests and try again.

HTTP Status Code: 429

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

CreatePaymentInstrument

Create a new payment instrument for a connector.

Request Syntax

```
POST /payments/createPaymentInstrument HTTP/1.1
X-Amzn-Bedrock-AgentCore-Payments-User-Id: userId
X-Amzn-Bedrock-AgentCore-Payments-Agent-Name: agentName
Content-type: application/json

{
  "clientToken": "string",
  "paymentConnectorId": "string",
  "paymentInstrumentDetails": { ... },
  "paymentInstrumentType": "string",
  "paymentManagerArn": "string"
}
```

URI Request Parameters

The request uses the following URI parameters.

[agentName](#)

The agent name associated with this request, used for observability.

Length Constraints: Minimum length of 0. Maximum length of 256.

[userId](#)

The user ID associated with this payment instrument.

Length Constraints: Minimum length of 0. Maximum length of 120.

Request Body

The request accepts the following data in JSON format.

[clientToken](#)

A unique, case-sensitive identifier that you provide to ensure the idempotency of the request.

Type: String

Length Constraints: Minimum length of 33. Maximum length of 256.

Pattern: `[a-zA-Z0-9](-*[a-zA-Z0-9]){0,256}`

Required: No

[paymentConnectorId](#)

The ID of the payment connector to use for this instrument.

Type: String

Length Constraints: Minimum length of 12. Maximum length of 211.

Pattern: `([0-9a-z][-]?){1,100}-[0-9a-z]{10}`

Required: Yes

[paymentInstrumentDetails](#)

The details of the payment instrument.

Type: [PaymentInstrumentDetails](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: Yes

[paymentInstrumentType](#)

The type of payment instrument being created.

Type: String

Valid Values: EMBEDDED_CRYPTOWALLET

Required: Yes

[paymentManagerArn](#)

The ARN of the payment manager that owns this payment instrument.

Type: String

Length Constraints: Minimum length of 66. Maximum length of 2048.

Pattern: `arn:(aws|aws-[a-z0-9-]+):bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:payment-manager/[a-z0-9]([a-z0-9-]{0,47}[a-z0-9])?-[a-z0-9]{10}`

Required: Yes

Response Syntax

```
HTTP/1.1 201
Content-type: application/json

{
  "paymentInstrument": {
    "createdAt": "string",
    "paymentConnectorId": "string",
    "paymentInstrumentDetails": { ... },
    "paymentInstrumentId": "string",
    "paymentInstrumentType": "string",
    "paymentManagerArn": "string",
    "status": "string",
    "updatedAt": "string",
    "userId": "string"
  }
}
```

Response Elements

If the action is successful, the service sends back an HTTP 201 response.

The following data is returned in JSON format by the service.

[paymentInstrument](#)

The created payment instrument.

Type: [PaymentInstrument](#) object

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

ConflictException

The exception that occurs when the request conflicts with the current state of the resource. This can happen when trying to modify a resource that is currently being modified by another request, or when trying to create a resource that already exists.

HTTP Status Code: 409

InternalServerError

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ServiceQuotaExceededException

The exception that occurs when the request would cause a service quota to be exceeded. Review your service quotas and either reduce your request rate or request a quota increase.

HTTP Status Code: 402

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

CreatePaymentSession

Create a new payment session.

Request Syntax

```
POST /payments/createPaymentSession HTTP/1.1
X-Amzn-Bedrock-AgentCore-Payments-User-Id: userId
X-Amzn-Bedrock-AgentCore-Payments-Agent-Name: agentName
Content-type: application/json
```

```
{
  "clientToken": "string",
  "expiryTimeInMinutes": number,
  "limits": {
    "maxSpendAmount": {
      "currency": "string",
      "value": "string"
    }
  },
  "paymentManagerArn": "string"
}
```

URI Request Parameters

The request uses the following URI parameters.

agentName

The agent name associated with this request, used for observability.

Length Constraints: Minimum length of 0. Maximum length of 256.

userId

The user ID associated with this payment session.

Length Constraints: Minimum length of 0. Maximum length of 120.

Request Body

The request accepts the following data in JSON format.

clientToken

A unique, case-sensitive identifier that you provide to ensure the idempotency of the request.

Type: String

Length Constraints: Minimum length of 33. Maximum length of 256.

Pattern: `[a-zA-Z0-9](-*[a-zA-Z0-9]){0,256}`

Required: No

expiryTimeInMinutes

The session expiry time in minutes. Must be between 15 and 480 minutes.

Type: Integer

Valid Range: Minimum value of 15. Maximum value of 480.

Required: Yes

limits

The spending limits for this payment session.

Type: [SessionLimits](#) object

Required: No

paymentManagerArn

The ARN of the payment manager that owns this session.

Type: String

Length Constraints: Minimum length of 66. Maximum length of 2048.

Pattern: `arn:(aws|aws-[a-z0-9-]+):bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:payment-manager/[a-z0-9]([a-z0-9-]{0,47}[a-z0-9])?-[a-z0-9]{10}`

Required: Yes

Response Syntax

```
HTTP/1.1 201
```

```
Content-type: application/json

{
  "paymentSession": {
    "availableLimits": {
      "availableSpendAmount": {
        "currency": "string",
        "value": "string"
      },
      "updatedAt": "string"
    },
    "createdAt": "string",
    "expiryTimeInMinutes": number,
    "limits": {
      "maxSpendAmount": {
        "currency": "string",
        "value": "string"
      }
    },
    "paymentManagerArn": "string",
    "paymentSessionId": "string",
    "updatedAt": "string",
    "userId": "string"
  }
}
```

Response Elements

If the action is successful, the service sends back an HTTP 201 response.

The following data is returned in JSON format by the service.

paymentSession

The created payment session.

Type: [PaymentSession](#) object

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

ConflictException

The exception that occurs when the request conflicts with the current state of the resource. This can happen when trying to modify a resource that is currently being modified by another request, or when trying to create a resource that already exists.

HTTP Status Code: 409

InternalServerError

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ServiceQuotaExceededException

The exception that occurs when the request would cause a service quota to be exceeded. Review your service quotas and either reduce your request rate or request a quota increase.

HTTP Status Code: 402

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DeleteABTest

Deletes an A/B test and its associated gateway rules.

Request Syntax

```
DELETE /ab-tests/abTestId HTTP/1.1
```

URI Request Parameters

The request uses the following URI parameters.

abTestId

The unique identifier of the A/B test to delete.

Pattern: `[a-zA-Z][a-zA-Z0-9-_{0,99}-[a-zA-Z0-9]{10}`

Required: Yes

Request Body

The request does not have a request body.

Response Syntax

```
HTTP/1.1 202  
Content-type: application/json
```

```
{  
  "abTestArn": "string",  
  "abTestId": "string",  
  "status": "string"  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 202 response.

The following data is returned in JSON format by the service.

abTestArn

The Amazon Resource Name (ARN) of the deleted A/B test.

Type: String

Pattern: `arn:aws[a-zA-Z-]*:bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:ab-test/[a-zA-Z][a-zA-Z0-9-]{0,99}-[a-zA-Z0-9]{10}`

abTestId

The unique identifier of the deleted A/B test.

Type: String

Pattern: `[a-zA-Z][a-zA-Z0-9-]{0,99}-[a-zA-Z0-9]{10}`

status

The status of the A/B test deletion operation.

Type: String

Valid Values: CREATING | ACTIVE | CREATE_FAILED | UPDATING | UPDATE_FAILED | DELETING | DELETE_FAILED | FAILED

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

ConflictException

The exception that occurs when the request conflicts with the current state of the resource. This can happen when trying to modify a resource that is currently being modified by another request, or when trying to create a resource that already exists.

HTTP Status Code: 409

InternalServerErrorException

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

UnauthorizedException

This exception is thrown when the JWT bearer token is invalid or not found for OAuth bearer token based access

HTTP Status Code: 401

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DeleteBatchEvaluation

Deletes a batch evaluation and its associated results.

Request Syntax

```
DELETE /evaluations/batch-evaluate/batchEvaluationId HTTP/1.1
```

URI Request Parameters

The request uses the following URI parameters.

batchEvaluationId

The unique identifier of the batch evaluation to delete.

Pattern: `[a-zA-Z][a-zA-Z0-9-_{0,99}-[a-zA-Z0-9]{10}`

Required: Yes

Request Body

The request does not have a request body.

Response Syntax

```
HTTP/1.1 202
```

```
Content-type: application/json
```

```
{  
  "batchEvaluationArn": "string",  
  "batchEvaluationId": "string",  
  "status": "string"  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 202 response.

The following data is returned in JSON format by the service.

batchEvaluationArn

The Amazon Resource Name (ARN) of the deleted batch evaluation.

Type: String

batchEvaluationId

The unique identifier of the deleted batch evaluation.

Type: String

Pattern: `[a-zA-Z][a-zA-Z0-9-_{0,99}-[a-zA-Z0-9]{10}`

status

The status of the batch evaluation deletion operation.

Type: String

Valid Values: PENDING | IN_PROGRESS | COMPLETED | COMPLETED_WITH_ERRORS | FAILED | STOPPING | STOPPED | DELETING

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

ConflictException

The exception that occurs when the request conflicts with the current state of the resource. This can happen when trying to modify a resource that is currently being modified by another request, or when trying to create a resource that already exists.

HTTP Status Code: 409

InternalServerErrorException

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

UnauthorizedException

This exception is thrown when the JWT bearer token is invalid or not found for OAuth bearer token based access

HTTP Status Code: 401

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)

- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DeleteEvent

Deletes an event from an AgentCore Memory resource. When you delete an event, it is permanently removed.

To use this operation, you must have the `bedrock-agentcore:DeleteEvent` permission.

Request Syntax

```
DELETE /memories/memoryId/actor/actorId/sessions/sessionId/events/eventId HTTP/1.1
```

URI Request Parameters

The request uses the following URI parameters.

actorId

The identifier of the actor associated with the event to delete.

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[a-zA-Z0-9][a-zA-Z0-9-_/]*(?::[a-zA-Z0-9-_/]+)*[a-zA-Z0-9-_/]*`

Required: Yes

eventId

The identifier of the event to delete.

Pattern: `[0-9]+#[a-zA-Z0-9-_/]*`

Required: Yes

memoryId

The identifier of the AgentCore Memory resource from which to delete the event.

Length Constraints: Minimum length of 12.

Pattern: `(arn:(aws|aws-cn|aws-us-gov):bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:memory/)?[a-zA-Z][a-zA-Z0-9-_/]{0,99}-[a-zA-Z0-9]{10}`

Required: Yes

sessionId

The identifier of the session containing the event to delete.

Length Constraints: Minimum length of 1. Maximum length of 100.

Pattern: `[a-zA-Z0-9][a-zA-Z0-9-_*]`*

Required: Yes

Request Body

The request does not have a request body.

Response Syntax

```
HTTP/1.1 200
Content-type: application/json

{
  "eventId": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

eventId

The identifier of the event that was deleted.

Type: String

Pattern: `[0-9]+#[a-zA-F0-9]+`

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

InvalidInputException

The input fails to satisfy the constraints specified by AgentCore. Check your input values and try again.

HTTP Status Code: 400

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ServiceException

The service encountered an internal error. Try your request again later.

HTTP Status Code: 500

ServiceQuotaExceededException

The exception that occurs when the request would cause a service quota to be exceeded. Review your service quotas and either reduce your request rate or request a quota increase.

HTTP Status Code: 402

ThrottledException

The request was denied due to request throttling. Reduce the frequency of requests and try again.

HTTP Status Code: 429

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DeleteMemoryRecord

Deletes a memory record from an AgentCore Memory resource. When you delete a memory record, it is permanently removed.

To use this operation, you must have the `bedrock-agentcore:DeleteMemoryRecord` permission.

Request Syntax

```
DELETE /memories/memoryId/memoryRecords/memoryRecordId HTTP/1.1
```

URI Request Parameters

The request uses the following URI parameters.

memoryId

The identifier of the AgentCore Memory resource from which to delete the memory record.

Length Constraints: Minimum length of 12.

Pattern: `(arn:(aws|aws-cn|aws-us-gov):bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:memory/)?[a-zA-Z][a-zA-Z0-9-_{0,99}]-[a-zA-Z0-9]{10}`

Required: Yes

memoryRecordId

The identifier of the memory record to delete.

Length Constraints: Minimum length of 40. Maximum length of 50.

Pattern: `mem-[a-zA-Z0-9-_{0,99}]*`

Required: Yes

Request Body

The request does not have a request body.

Response Syntax

```
HTTP/1.1 200
Content-type: application/json

{
  "memoryRecordId": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

[memoryRecordId](#)

The identifier of the memory record that was deleted.

Type: String

Length Constraints: Minimum length of 40. Maximum length of 50.

Pattern: mem-[a-zA-Z0-9-_*]

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

InvalidInputException

The input fails to satisfy the constraints specified by AgentCore. Check your input values and try again.

HTTP Status Code: 400

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ServiceException

The service encountered an internal error. Try your request again later.

HTTP Status Code: 500

ServiceQuotaExceededException

The exception that occurs when the request would cause a service quota to be exceeded. Review your service quotas and either reduce your request rate or request a quota increase.

HTTP Status Code: 402

ThrottledException

The request was denied due to request throttling. Reduce the frequency of requests and try again.

HTTP Status Code: 429

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)

- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DeletePaymentInstrument

Deletes a payment instrument. This is a soft delete operation that preserves the record for audit and compliance purposes.

Request Syntax

```
POST /payments/deletePaymentInstrument HTTP/1.1
X-Amzn-Bedrock-AgentCore-Payments-User-Id: userId
Content-type: application/json

{
  "paymentConnectorId": "string",
  "paymentInstrumentId": "string",
  "paymentManagerArn": "string"
}
```

URI Request Parameters

The request uses the following URI parameters.

userId

The user ID making the delete request. Must match the instrument's userId.

Length Constraints: Minimum length of 0. Maximum length of 120.

Request Body

The request accepts the following data in JSON format.

paymentConnectorId

The payment connector ID. Must match the instrument's paymentConnectorId.

Type: String

Length Constraints: Minimum length of 12. Maximum length of 211.

Pattern: ($[0-9a-z][-]?$){1,100}- $[0-9a-z]$ {10}

Required: Yes

paymentInstrumentId

The payment instrument ID to delete.

Type: String

Length Constraints: Fixed length of 34.

Pattern: `payment-instrument-[0-9a-zA-Z-]{15}`

Required: Yes

paymentManagerArn

The payment manager ARN. Must match the instrument's `paymentManagerArn`.

Type: String

Length Constraints: Minimum length of 66. Maximum length of 2048.

Pattern: `arn:(aws|aws-[a-z0-9-]+):bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:payment-manager/[a-z0-9]([a-z0-9-]{0,47}[a-z0-9])?-[a-z0-9]{10}`

Required: Yes

Response Syntax

```
HTTP/1.1 200
Content-type: application/json
```

```
{
  "status": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

status

The status of the instrument after deletion. Always DELETED for successful soft delete.

Type: String

Valid Values: INITIATED | ACTIVE | FAILED | DELETED

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

InternalServerError

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DeletePaymentSession

Deletes a payment session. This permanently removes the payment session record.

Request Syntax

```
POST /payments/deletePaymentSession HTTP/1.1
X-Amzn-Bedrock-AgentCore-Payments-User-Id: userId
Content-type: application/json
```

```
{
  "paymentManagerArn": "string",
  "paymentSessionId": "string"
}
```

URI Request Parameters

The request uses the following URI parameters.

userId

The user ID making the delete request. Must match the session's `userId`.

Length Constraints: Minimum length of 0. Maximum length of 120.

Request Body

The request accepts the following data in JSON format.

paymentManagerArn

The payment manager ARN. Must match the session's `paymentManagerArn`.

Type: String

Length Constraints: Minimum length of 66. Maximum length of 2048.

Pattern: `arn:(aws|aws-[a-z0-9-]+):bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:payment-manager/[a-z0-9]([a-z0-9-]{0,47}[a-z0-9])?-[a-z0-9]{10}`

Required: Yes

paymentSessionId

The payment session ID to delete.

Type: String

Length Constraints: Fixed length of 31.

Pattern: payment-session-[0-9a-zA-Z-]{15}

Required: Yes

Response Syntax

```
HTTP/1.1 200
Content-type: application/json
```

```
{
  "status": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

status

The status of the deletion. Always DELETED for successful hard delete.

Type: String

Valid Values: ACTIVE | EXPIRED | DELETED

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

InternalServerError

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DeleteRecommendation

Deletes a recommendation and its associated results.

Request Syntax

```
DELETE /recommendations/recommendationId HTTP/1.1
```

URI Request Parameters

The request uses the following URI parameters.

recommendationId

The unique identifier of the recommendation to delete.

Pattern: `[0-9a-zA-Z_-]{1,48}-[0-9A-Z]{10}`

Required: Yes

Request Body

The request does not have a request body.

Response Syntax

```
HTTP/1.1 202
Content-type: application/json

{
  "recommendationId": "string",
  "status": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 202 response.

The following data is returned in JSON format by the service.

recommendationId

The unique identifier of the deleted recommendation.

Type: String

Pattern: `[0-9a-zA-Z_-]{1,48}-[0-9A-Z]{10}`

status

The status of the recommendation deletion operation.

Type: String

Valid Values: PENDING | IN_PROGRESS | COMPLETED | FAILED | DELETING

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

ConflictException

The exception that occurs when the request conflicts with the current state of the resource. This can happen when trying to modify a resource that is currently being modified by another request, or when trying to create a resource that already exists.

HTTP Status Code: 409

InternalServerError

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

Evaluate

Performs on-demand evaluation of agent traces using a specified evaluator. This synchronous API accepts traces in OpenTelemetry format and returns immediate scoring results with detailed explanations.

Request Syntax

```
POST /evaluations/evaluate/evaluatorId HTTP/1.1
```

```
Content-type: application/json
```

```
{
  "evaluationInput": { ... },
  "evaluationReferenceInputs": [
    {
      "assertions": [
        { ... }
      ],
      "context": { ... },
      "expectedResponse": { ... },
      "expectedTrajectory": {
        "toolNames": [ "string" ]
      }
    }
  ],
  "evaluationTarget": { ... }
}
```

URI Request Parameters

The request uses the following URI parameters.

evaluatorId

The unique identifier of the evaluator to use for scoring. Can be a built-in evaluator (e.g., `Builtin.Helpfulness`, `Builtin.Correctness`) or a custom evaluator Id created through the control plane API.

Pattern: `(Builtin.[a-zA-Z0-9_-]+|[a-zA-Z][a-zA-Z0-9-_{0,99}}-[a-zA-Z0-9]{10})`

Required: Yes

Request Body

The request accepts the following data in JSON format.

evaluationInput

The input data containing agent session spans to be evaluated. Includes a list of spans in OpenTelemetry format from supported frameworks like Strands (AgentCore Runtime) or LangGraph with OpenInference instrumentation.

Type: [EvaluationInput](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: Yes

evaluationReferenceInputs

Ground truth data to compare against agent responses during evaluation. Allows to provide expected responses, assertions, and expected tool trajectories at different evaluation levels. Session-level reference inputs apply to the entire conversation, while trace-level reference inputs target specific request-response interactions identified by trace ID.

Type: Array of [EvaluationReferenceInput](#) objects

Array Members: Minimum number of 1 item. Maximum number of 1000 items.

Required: No

evaluationTarget

The specific trace or span IDs to evaluate within the provided input. Allows targeting evaluation at different levels: individual tool calls, single request-response interactions (traces), or entire conversation sessions.

Type: [EvaluationTarget](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: No

Response Syntax

```
HTTP/1.1 200
```

```
Content-type: application/json

{
  "evaluationResults": [
    {
      "context": { ... },
      "errorCode": "string",
      "errorMessage": "string",
      "evaluatorArn": "string",
      "evaluatorId": "string",
      "evaluatorName": "string",
      "explanation": "string",
      "ignoredReferenceInputFields": [ "string" ],
      "label": "string",
      "tokenUsage": {
        "inputTokens": number,
        "outputTokens": number,
        "totalTokens": number
      },
      "value": number
    }
  ]
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

evaluationResults

The detailed evaluation results containing scores, explanations, and metadata. Includes the evaluator information, numerical or categorical ratings based on the evaluator's rating scale, and token usage statistics for the evaluation process.

Type: Array of [EvaluationResultContent](#) objects

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

ConflictException

The exception that occurs when the request conflicts with the current state of the resource. This can happen when trying to modify a resource that is currently being modified by another request, or when trying to create a resource that already exists.

HTTP Status Code: 409

DuplicateIdException

An exception thrown when attempting to create a resource with an identifier that already exists.

HTTP Status Code: 409

InternalServerError

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ServiceQuotaExceededException

The exception that occurs when the request would cause a service quota to be exceeded. Review your service quotas and either reduce your request rate or request a quota increase.

HTTP Status Code: 402

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

UnauthorizedException

This exception is thrown when the JWT bearer token is invalid or not found for OAuth bearer token based access

HTTP Status Code: 401

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetABTest

Retrieves detailed information about an A/B test, including its configuration, status, and statistical results.

Request Syntax

```
GET /ab-tests/abTestId HTTP/1.1
```

URI Request Parameters

The request uses the following URI parameters.

[abTestId](#)

The unique identifier of the A/B test to retrieve.

Pattern: `[a-zA-Z][a-zA-Z0-9-_{0,99}-[a-zA-Z0-9]{10}`

Required: Yes

Request Body

The request does not have a request body.

Response Syntax

```
HTTP/1.1 200
```

```
Content-type: application/json
```

```
{
  "abTestArn": "string",
  "abTestId": "string",
  "createdAt": number,
  "currentRunId": "string",
  "description": "string",
  "errorDetails": [ "string" ],
  "evaluationConfig": { ... },
  "executionStatus": "string",
  "gatewayArn": "string",
  "gatewayFilter": {
```

```
    "targetPaths": [ "string" ]
  },
  "maxDurationExpiresAt": number,
  "name": "string",
  "results": {
    "analysisTimestamp": number,
    "evaluatorMetrics": [
      {
        "controlStats": {
          "mean": number,
          "sampleSize": number,
          "variantName": "string"
        },
        "evaluatorArn": "string",
        "variantResults": [
          {
            "absoluteChange": number,
            "confidenceInterval": {
              "lower": number,
              "upper": number
            },
            "isSignificant": boolean,
            "mean": number,
            "percentChange": number,
            "pValue": number,
            "sampleSize": number,
            "variantName": "string"
          }
        ]
      }
    ]
  },
  "roleArn": "string",
  "startedAt": number,
  "status": "string",
  "stoppedAt": number,
  "updatedAt": number,
  "variants": [
    {
      "name": "string",
      "variantConfiguration": {
        "configurationBundle": {
          "bundleArn": "string",
          "bundleVersion": "string"
        }
      }
    }
  ]
}
```

```
    },
    "target": {
      "name": "string"
    }
  },
  "weight": number
}
]
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

[abTestArn](#)

The Amazon Resource Name (ARN) of the A/B test.

Type: String

Pattern: `arn:aws[a-zA-Z-]*:bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:ab-test/[a-zA-Z][a-zA-Z0-9-]{0,99}-[a-zA-Z0-9]{10}`

[abTestId](#)

The unique identifier of the A/B test.

Type: String

Pattern: `[a-zA-Z][a-zA-Z0-9-]{0,99}-[a-zA-Z0-9]{10}`

[createdAt](#)

The timestamp when the A/B test was created.

Type: Timestamp

[currentRunId](#)

The identifier of the current run of the A/B test.

Type: String

description

The description of the A/B test.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 200.

errorDetails

The error details if the A/B test encountered failures.

Type: Array of strings

Array Members: Minimum number of 0 items. Maximum number of 1 item.

Length Constraints: Minimum length of 0. Maximum length of 1000.

evaluationConfig

The evaluation configuration for measuring variant performance.

Type: [ABTestEvaluationConfig](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

executionStatus

The execution status indicating whether the A/B test is currently running.

Type: String

Valid Values: PAUSED | RUNNING | STOPPED | NOT_STARTED

gatewayArn

The Amazon Resource Name (ARN) of the gateway used for traffic splitting.

Type: String

Pattern: `arn:aws(|-cn|-us-gov):bedrock-agentcore:[a-z0-9-]{1,20}:[0-9]{12}:gateway/([0-9a-z][-]?)^{1,48}-[a-z0-9]{10}`

gatewayFilter

The gateway filter restricting which target paths are included.

Type: [GatewayFilter](#) object

[maxDurationExpiresAt](#)

The timestamp when the A/B test will automatically expire.

Type: Timestamp

[name](#)

The name of the A/B test.

Type: String

Pattern: `[a-zA-Z][a-zA-Z0-9_]{0,47}`

[results](#)

The statistical results of the A/B test, including per-evaluator metrics and significance analysis.

Type: [ABTestResults](#) object

[roleArn](#)

The IAM role ARN used by the A/B test.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: `arn:aws(-[^:]+)?:iam::([0-9]{12})?:role/.+`

[startedAt](#)

The timestamp when the A/B test was started.

Type: Timestamp

[status](#)

The current status of the A/B test.

Type: String

Valid Values: `CREATING | ACTIVE | CREATE_FAILED | UPDATING | UPDATE_FAILED | DELETING | DELETE_FAILED | FAILED`

stoppedAt

The timestamp when the A/B test was stopped.

Type: Timestamp

updatedAt

The timestamp when the A/B test was last updated.

Type: Timestamp

variants

The list of variants in the A/B test.

Type: Array of [Variant](#) objects

Array Members: Fixed number of 2 items.

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

InternalServerErrorException

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

UnauthorizedException

This exception is thrown when the JWT bearer token is invalid or not found for OAuth bearer token based access

HTTP Status Code: 401

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)

- [AWS SDK for Ruby V3](#)

GetAgentCard

Retrieves the A2A agent card associated with an AgentCore Runtime agent.

Request Syntax

```
GET /runtimes/agentRuntimeArn/invocations/.well-known/agent-card.json?  
qualifier=qualifier HTTP/1.1  
X-Amzn-Bedrock-AgentCore-Runtime-Session-Id: runtimeSessionId
```

URI Request Parameters

The request uses the following URI parameters.

agentRuntimeArn

The ARN of the AgentCore Runtime agent for which you want to get the A2A agent card.

Required: Yes

qualifier

Optional qualifier to specify an agent alias, such as `prodcode>` or `dev`. If you don't provide a value, the DEFAULT alias is used.

runtimeSessionId

The session ID that the AgentCore Runtime agent is using.

Length Constraints: Minimum length of 33. Maximum length of 256.

Request Body

The request does not have a request body.

Response Syntax

```
HTTP/1.1 statusCode  
X-Amzn-Bedrock-AgentCore-Runtime-Session-Id: runtimeSessionId
```

Response Elements

If the action is successful, the service sends back the following HTTP response.

statusCode

The status code of the request.

The response returns the following HTTP headers.

runtimeSessionId

The ID of the session associated with the AgentCore Runtime agent.

Length Constraints: Minimum length of 1. Maximum length of 100.

Pattern: `[a-zA-Z0-9][a-zA-Z0-9-_*]`*

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

InternalServerError

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

RuntimeClientError

The exception that occurs when there is an error in the runtime client. This can happen due to network issues, invalid configuration, or other client-side problems. Check the error message for specific details about the error.

HTTP Status Code: 424

ServiceQuotaExceededException

The exception that occurs when the request would cause a service quota to be exceeded. Review your service quotas and either reduce your request rate or request a quota increase.

HTTP Status Code: 402

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)

- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetBatchEvaluation

Retrieves detailed information about a batch evaluation, including its status, configuration, results, and any error details.

Request Syntax

```
GET /evaluations/batch-evaluate/batchEvaluationId HTTP/1.1
```

URI Request Parameters

The request uses the following URI parameters.

batchEvaluationId

The unique identifier of the batch evaluation to retrieve.

Pattern: `[a-zA-Z][a-zA-Z0-9-_{0,99}-[a-zA-Z0-9]{10}`

Required: Yes

Request Body

The request does not have a request body.

Response Syntax

```
HTTP/1.1 200
```

```
Content-type: application/json
```

```
{
  "batchEvaluationArn": "string",
  "batchEvaluationId": "string",
  "batchEvaluationName": "string",
  "createdAt": "string",
  "dataSourceConfig": { ... },
  "description": "string",
  "errorDetails": [ "string" ],
  "evaluationResults": {
    "evaluatorSummaries": [
      {
```

```
    "evaluatorId": "string",
    "statistics": {
      "averageScore": number
    },
    "totalEvaluated": number,
    "totalFailed": number
  }
],
"numberOfSessionsCompleted": number,
"numberOfSessionsFailed": number,
"numberOfSessionsIgnored": number,
"numberOfSessionsInProgress": number,
"totalNumberOfSessions": number
},
"evaluators": [
  {
    "evaluatorId": "string"
  }
],
"executionSummaryResult": {
  "executionSummaries": [
    {
      "affectedSessionCount": number,
      "affectedSessions": [
        {
          "approachTaken": "string",
          "finalOutcome": "string",
          "sessionId": "string"
        }
      ],
      "clusterId": number,
      "description": "string",
      "name": "string"
    }
  ]
},
"failureAnalysisResult": {
  "failures": [
    {
      "affectedSessionCount": number,
      "clusterId": number,
      "description": "string",
      "name": "string",
      "subCategories": [
```

```

    {
      "affectedSessionCount": number,
      "clusterId": number,
      "description": "string",
      "name": "string",
      "rootCauses": [
        {
          "affectedSessionCount": number,
          "affectedSessions": [
            {
              "explanation": "string",
              "failureSpans": [
                {
                  "signals": [
                    {
                      "category": "string",
                      "confidence": number,
                      "evidence": "string"
                    }
                  ],
                  "spanId": "string",
                  "traceId": "string"
                }
              ],
              "fixType": "string",
              "recommendation": "string",
              "sessionId": "string"
            }
          ],
          "clusterId": number,
          "name": "string",
          "recommendation": "string",
          "rootCause": "string"
        }
      ]
    }
  ],
  "insights": [
    {
      "insightId": "string"
    }
  ]
},

```

```

],
  "kmsKeyArn": "string",
  "outputConfig": { ... },
  "status": "string",
  "updatedAt": "string",
  "userIntentResult": {
    "userIntents": [
      {
        "affectedSessionCount": number,
        "affectedSessions": [
          {
            "sessionId": "string",
            "userMessages": [ "string" ]
          }
        ],
        "clusterId": number,
        "description": "string",
        "name": "string"
      }
    ]
  }
}

```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

batchEvaluationArn

The Amazon Resource Name (ARN) of the batch evaluation.

Type: String

batchEvaluationId

The unique identifier of the batch evaluation.

Type: String

Pattern: `[a-zA-Z][a-zA-Z0-9-_{0,99}-[a-zA-Z0-9]{10}`

batchEvaluationName

The name of the batch evaluation.

Type: String

Pattern: `[a-zA-Z][a-zA-Z0-9_]{0,47}`

createdAt

The timestamp when the batch evaluation was created.

Type: Timestamp

dataSourceConfig

The data source configuration specifying where agent traces are pulled from.

Type: [DataSourceConfig](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

description

The description of the batch evaluation.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 200.

errorDetails

The error details if the batch evaluation encountered failures.

Type: Array of strings

Array Members: Minimum number of 0 items. Maximum number of 1 item.

Length Constraints: Minimum length of 0. Maximum length of 1000.

evaluationResults

The aggregated evaluation results, including session completion counts and evaluator score summaries.

Type: [EvaluationJobResults](#) object

evaluators

The list of evaluators applied during the batch evaluation.

Type: Array of [Evaluator](#) objects

executionSummaryResult

The execution summary clustering results from insights, containing grouped execution patterns across evaluated sessions.

Type: [ExecutionSummaryClusteringResultContent](#) object

failureAnalysisResult

The failure analysis results from insights, containing categorized failure clusters with root causes and recommendations.

Type: [FailureAnalysisResultContent](#) object

insights

The list of insight analyses applied during the batch evaluation.

Type: Array of [Insight](#) objects

Array Members: Minimum number of 0 items. Maximum number of 10 items.

kmsKeyArn

The ARN of the AWS KMS key used to encrypt evaluation data.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: `arn:aws(|-cn|-us-gov):kms:[a-zA-Z0-9-]*:[0-9]{12}:key/[a-zA-Z0-9-]{36}`

outputConfig

The output configuration specifying where evaluation results are written.

Type: [OutputConfig](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

status

The current status of the batch evaluation.

Type: String

Valid Values: PENDING | IN_PROGRESS | COMPLETED | COMPLETED_WITH_ERRORS | FAILED | STOPPING | STOPPED | DELETING

updatedAt

The timestamp when the batch evaluation was last updated.

Type: Timestamp

userIntentResult

The user intent clustering results from insights, containing grouped user intents across evaluated sessions.

Type: [UserIntentClusteringResultContent](#) object

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

InternalServerError

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

UnauthorizedException

This exception is thrown when the JWT bearer token is invalid or not found for OAuth bearer token based access

HTTP Status Code: 401

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)

- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetBrowserSession

Retrieves detailed information about a specific browser session in Amazon Bedrock AgentCore. This operation returns the session's configuration, current status, associated streams, and metadata.

To get a browser session, you must specify both the browser identifier and the session ID. The response includes information about the session's viewport configuration, timeout settings, and stream endpoints.

The following operations are related to `GetBrowserSession`:

- [StartBrowserSession](#)
- [ListBrowserSessions](#)
- [StopBrowserSession](#)

Request Syntax

```
GET /browsers/browserIdentifier/sessions/get?sessionId=sessionId HTTP/1.1
```

URI Request Parameters

The request uses the following URI parameters.

[browserIdentifier](#)

The unique identifier of the browser associated with the session.

Required: Yes

[sessionId](#)

The unique identifier of the browser session to retrieve.

Pattern: `[0-9a-zA-Z]{1,40}`

Required: Yes

Request Body

The request does not have a request body.

Response Syntax

HTTP/1.1 200

Content-type: application/json

```
{
  "browserIdentifier": "string",
  "certificates": [
    {
      "location": { ... }
    }
  ],
  "createdAt": "string",
  "enterprisePolicies": [
    {
      "location": { ... },
      "type": "string"
    }
  ],
  "extensions": [
    {
      "location": { ... }
    }
  ],
  "lastUpdatedAt": "string",
  "name": "string",
  "profileConfiguration": {
    "profileIdentifier": "string"
  },
  "proxyConfiguration": {
    "bypass": {
      "domainPatterns": [ "string" ]
    },
    "proxies": [
      { ... }
    ]
  },
  "sessionId": "string",
  "sessionReplayArtifact": "string",
  "sessionTimeoutSeconds": number,
  "status": "string",
  "streams": {
    "automationStream": {
```

```
    "streamEndpoint": "string",
    "streamStatus": "string"
  },
  "liveViewStream": {
    "streamEndpoint": "string"
  }
},
"viewport": {
  "height": number,
  "width": number
}
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

browserIdentifier

The identifier of the browser.

Type: String

certificates

The list of certificates installed in the browser session.

Type: Array of [Certificate](#) objects

Array Members: Minimum number of 1 item. Maximum number of 200 items.

createdAt

The time at which the browser session was created.

Type: Timestamp

enterprisePolicies

A list of files containing enterprise policies for the browser session.

Type: Array of [BrowserEnterprisePolicy](#) objects

Array Members: Minimum number of 0 items. Maximum number of 100 items.

extensions

The list of browser extensions that are configured in the browser session.

Type: Array of [BrowserExtension](#) objects

Array Members: Minimum number of 1 item. Maximum number of 10 items.

lastUpdatedAt

The time at which the browser session was last updated.

Type: Timestamp

name

The name of the browser session.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 100.

profileConfiguration

The browser profile configuration associated with this session. Contains the profile identifier that links to persistent browser data such as cookies and local storage.

Type: [BrowserProfileConfiguration](#) object

proxyConfiguration

The active proxy configuration for this browser session. This field is only present if proxy configuration was provided when the session was started using `StartBrowserSession`. The configuration includes proxy servers, domain bypass rules and the proxy authentication credentials.

Type: [ProxyConfiguration](#) object

sessionId

The identifier of the browser session.

Type: String

Pattern: `[0-9a-zA-Z]{1,40}`

[sessionReplayArtifact](#)

The artifact containing the session replay information.

Type: String

[sessionTimeoutSeconds](#)

The timeout period for the browser session in seconds.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 28800.

[status](#)

The current status of the browser session. Possible values include ACTIVE, STOPPING, and STOPPED.

Type: String

Valid Values: READY | TERMINATED

[streams](#)

The streams associated with this browser session. These include the automation stream and live view stream.

Type: [BrowserSessionStream](#) object

[viewPort](#)

The configuration that defines the dimensions of a browser viewport in a browser session. The viewport determines the visible area of web content and affects how web pages are rendered and displayed. Proper viewport configuration ensures that web content is displayed correctly for the agent's browsing tasks.

Type: [ViewPort](#) object

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

InternalServerError

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetCodeInterpreterSession

Retrieves detailed information about a specific code interpreter session in Amazon Bedrock AgentCore. This operation returns the session's configuration, current status, and metadata.

To get a code interpreter session, you must specify both the code interpreter identifier and the session ID. The response includes information about the session's timeout settings and current status.

The following operations are related to `GetCodeInterpreterSession`:

- [StartCodeInterpreterSession](#)
- [ListCodeInterpreterSessions](#)
- [StopCodeInterpreterSession](#)

Request Syntax

```
GET /code-interpreters/codeInterpreterIdentifier/sessions/get?sessionId=sessionId
HTTP/1.1
```

URI Request Parameters

The request uses the following URI parameters.

[codeInterpreterIdentifier](#)

The unique identifier of the code interpreter associated with the session.

Required: Yes

[sessionId](#)

The unique identifier of the code interpreter session to retrieve.

Pattern: `[0-9a-zA-Z]{1,40}`

Required: Yes

Request Body

The request does not have a request body.

Response Syntax

```
HTTP/1.1 200
Content-type: application/json

{
  "certificates": [
    {
      "location": { ... }
    }
  ],
  "codeInterpreterIdentifier": "string",
  "createdAt": "string",
  "name": "string",
  "sessionId": "string",
  "sessionTimeoutSeconds": number,
  "status": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

certificates

The list of certificates installed in the code interpreter session.

Type: Array of [Certificate](#) objects

Array Members: Minimum number of 1 item. Maximum number of 200 items.

codeInterpreterIdentifier

The identifier of the code interpreter.

Type: String

createdAt

The time at which the code interpreter session was created.

Type: Timestamp

name

The name of the code interpreter session.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 100.

sessionId

The identifier of the code interpreter session.

Type: String

Pattern: `[0-9a-zA-Z]{1,40}`

sessionTimeoutSeconds

The timeout period for the code interpreter session in seconds.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 28800.

status

The current status of the code interpreter session. Possible values include ACTIVE, STOPPING, and STOPPED.

Type: String

Valid Values: READY | TERMINATED

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

InternalServerError

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetEvent

Retrieves information about a specific event in an AgentCore Memory resource.

To use this operation, you must have the `bedrock-agentcore:GetEvent` permission.

Request Syntax

```
GET /memories/memoryId/actor/actorId/sessions/sessionId/events/eventId HTTP/1.1
```

URI Request Parameters

The request uses the following URI parameters.

actorId

The identifier of the actor associated with the event.

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[a-zA-Z0-9][a-zA-Z0-9-_/]*(?::[a-zA-Z0-9-_/]+)*[a-zA-Z0-9-_/]*`

Required: Yes

eventId

The identifier of the event to retrieve.

Pattern: `[0-9]+#[a-fA-F0-9]+`

Required: Yes

memoryId

The identifier of the AgentCore Memory resource containing the event.

Length Constraints: Minimum length of 12.

Pattern: `(arn:(aws|aws-cn|aws-us-gov):bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:memory/)?[a-zA-Z][a-zA-Z0-9-_{0,99}]-[a-zA-Z0-9]{10}`

Required: Yes

sessionId

The identifier of the session containing the event.

Length Constraints: Minimum length of 1. Maximum length of 100.

Pattern: `[a-zA-Z0-9][a-zA-Z0-9-_*]*`

Required: Yes

Request Body

The request does not have a request body.

Response Syntax

```
HTTP/1.1 200
Content-type: application/json

{
  "event": {
    "actorId": "string",
    "branch": {
      "name": "string",
      "rootEventId": "string"
    },
    "eventId": "string",
    "eventTimestamp": number,
    "memoryId": "string",
    "metadata": {
      "string" : { ... }
    },
    "payload": [
      { ... }
    ],
    "sessionId": "string"
  }
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

event

The requested event information.

Type: [Event](#) object

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

InvalidInputException

The input fails to satisfy the constraints specified by AgentCore. Check your input values and try again.

HTTP Status Code: 400

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ServiceException

The service encountered an internal error. Try your request again later.

HTTP Status Code: 500

ServiceQuotaExceededException

The exception that occurs when the request would cause a service quota to be exceeded. Review your service quotas and either reduce your request rate or request a quota increase.

HTTP Status Code: 402

ThrottledException

The request was denied due to request throttling. Reduce the frequency of requests and try again.

HTTP Status Code: 429

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetMemoryRecord

Retrieves a specific memory record from an AgentCore Memory resource.

To use this operation, you must have the `bedrock-agentcore:GetMemoryRecord` permission.

Request Syntax

```
GET /memories/memoryId/memoryRecord/memoryRecordId HTTP/1.1
```

URI Request Parameters

The request uses the following URI parameters.

memoryId

The identifier of the AgentCore Memory resource containing the memory record.

Length Constraints: Minimum length of 12.

Pattern: `(arn:(aws|aws-cn|aws-us-gov):bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:memory/)?[a-zA-Z][a-zA-Z0-9-_{0,99}]-[a-zA-Z0-9]{10}`

Required: Yes

memoryRecordId

The identifier of the memory record to retrieve.

Length Constraints: Minimum length of 40. Maximum length of 50.

Pattern: `mem-[a-zA-Z0-9-_{0,99}]*`

Required: Yes

Request Body

The request does not have a request body.

Response Syntax

```
HTTP/1.1 200
```

```
Content-type: application/json

{
  "memoryRecord": {
    "content": { ... },
    "createdAt": number,
    "memoryRecordId": "string",
    "memoryStrategyId": "string",
    "metadata": {
      "string" : { ... }
    },
    "namespaces": [ "string" ]
  }
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

[memoryRecord](#)

The requested memory record.

Type: [MemoryRecord](#) object

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

InvalidInputException

The input fails to satisfy the constraints specified by AgentCore. Check your input values and try again.

HTTP Status Code: 400

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ServiceException

The service encountered an internal error. Try your request again later.

HTTP Status Code: 500

ServiceQuotaExceededException

The exception that occurs when the request would cause a service quota to be exceeded. Review your service quotas and either reduce your request rate or request a quota increase.

HTTP Status Code: 402

ThrottledException

The request was denied due to request throttling. Reduce the frequency of requests and try again.

HTTP Status Code: 429

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)

- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetPaymentInstrument

Get a payment instrument by ID.

Request Syntax

```
POST /payments/getPaymentInstrument HTTP/1.1
X-Amzn-Bedrock-AgentCore-Payments-User-Id: userId
X-Amzn-Bedrock-AgentCore-Payments-Agent-Name: agentName
Content-type: application/json
```

```
{
  "paymentConnectorId": "string",
  "paymentInstrumentId": "string",
  "paymentManagerArn": "string"
}
```

URI Request Parameters

The request uses the following URI parameters.

agentName

The agent name associated with this request, used for observability.

Length Constraints: Minimum length of 0. Maximum length of 256.

userId

The user ID associated with this payment instrument.

Length Constraints: Minimum length of 0. Maximum length of 120.

Request Body

The request accepts the following data in JSON format.

paymentConnectorId

The ID of the payment connector.

Type: String

Length Constraints: Minimum length of 12. Maximum length of 211.

Pattern: (`[0-9a-z][-]?{1,100}`)-`[0-9a-z]{10}`

Required: No

paymentInstrumentId

The ID of the payment instrument to retrieve.

Type: String

Length Constraints: Fixed length of 34.

Pattern: `payment-instrument-[0-9a-zA-Z-]{15}`

Required: Yes

paymentManagerArn

The ARN of the payment manager that owns this payment instrument.

Type: String

Length Constraints: Minimum length of 66. Maximum length of 2048.

Pattern: `arn:(aws|aws-[a-z0-9-]+):bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:payment-manager/[a-z0-9]([a-z0-9-]{0,47}[a-z0-9])?-[a-z0-9]{10}`

Required: Yes

Response Syntax

```
HTTP/1.1 200
```

```
Content-type: application/json
```

```
{
  "paymentInstrument": {
    "createdAt": "string",
    "paymentConnectorId": "string",
    "paymentInstrumentDetails": { ... },
    "paymentInstrumentId": "string",
    "paymentInstrumentType": "string",
    "paymentManagerArn": "string",
```

```
  "status": "string",  
  "updatedAt": "string",  
  "userId": "string"  
}  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

paymentInstrument

The payment instrument details.

Type: [PaymentInstrument](#) object

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

InternalServerError

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetPaymentInstrumentBalance

Get the balance of a payment instrument.

Request Syntax

```
POST /payments/getPaymentInstrumentBalance HTTP/1.1
X-Amzn-Bedrock-AgentCore-Payments-User-Id: userId
X-Amzn-Bedrock-AgentCore-Payments-Agent-Name: agentName
Content-type: application/json
```

```
{
  "chain": "string",
  "paymentConnectorId": "string",
  "paymentInstrumentId": "string",
  "paymentManagerArn": "string",
  "token": "string"
}
```

URI Request Parameters

The request uses the following URI parameters.

agentName

The agent name associated with this request, used for observability.

Length Constraints: Minimum length of 0. Maximum length of 256.

userId

The user ID associated with this payment instrument.

Length Constraints: Minimum length of 0. Maximum length of 120.

Request Body

The request accepts the following data in JSON format.

chain

The specific blockchain chain to query balance on. Required because balances are chain-specific.

Type: String

Valid Values: BASE | BASE_SEPOLIA | ETHEREUM | SOLANA | SOLANA_DEVNET

Required: Yes

paymentConnectorId

The ID of the payment connector associated with this instrument.

Type: String

Length Constraints: Minimum length of 12. Maximum length of 211.

Pattern: (`[0-9a-z][-]?{1,100}`)-`[0-9a-z]{10}`

Required: Yes

paymentInstrumentId

The ID of the payment instrument to query balance for.

Type: String

Length Constraints: Fixed length of 34.

Pattern: `payment-instrument-[0-9a-zA-Z-]{15}`

Required: Yes

paymentManagerArn

The ARN of the payment manager that owns this payment instrument.

Type: String

Length Constraints: Minimum length of 66. Maximum length of 2048.

Pattern: `arn:(aws|aws-[a-z0-9-]+):bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:payment-manager/[a-z0-9]([a-z0-9-]{0,47}[a-z0-9])?-[a-z0-9]{10}`

Required: Yes

token

The token to query balance for. Only tokens supported for X402 payments are returned.

Type: String

Valid Values: USDC

Required: Yes

Response Syntax

```
HTTP/1.1 200
Content-type: application/json

{
  "paymentInstrumentId": "string",
  "tokenBalance": {
    "amount": "string",
    "chain": "string",
    "decimals": number,
    "network": "string",
    "token": "string"
  }
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

paymentInstrumentId

The ID of the payment instrument.

Type: String

Length Constraints: Fixed length of 34.

Pattern: payment-instrument-[0-9a-zA-Z-]{15}

tokenBalance

The balance of the supported token on the requested chain.

Type: [TokenBalance](#) object

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

InternalServerError

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetPaymentSession

Get a payment session.

Request Syntax

```
POST /payments/getPaymentSession HTTP/1.1
X-Amzn-Bedrock-AgentCore-Payments-User-Id: userId
X-Amzn-Bedrock-AgentCore-Payments-Agent-Name: agentName
Content-type: application/json

{
  "paymentManagerArn": "string",
  "paymentSessionId": "string"
}
```

URI Request Parameters

The request uses the following URI parameters.

[agentName](#)

The agent name associated with this request, used for observability.

Length Constraints: Minimum length of 0. Maximum length of 256.

[userId](#)

The user ID associated with this payment session.

Length Constraints: Minimum length of 0. Maximum length of 120.

Request Body

The request accepts the following data in JSON format.

[paymentManagerArn](#)

The ARN of the payment manager that owns this session.

Type: String

Length Constraints: Minimum length of 66. Maximum length of 2048.

Pattern: `arn:(aws|aws-[a-z0-9-]+):bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:payment-manager/[a-z0-9]([a-z0-9-]{0,47}[a-z0-9])?-[a-z0-9]{10}`

Required: Yes

paymentSessionId

The ID of the payment session to retrieve.

Type: String

Length Constraints: Fixed length of 31.

Pattern: `payment-session-[0-9a-zA-Z-]{15}`

Required: Yes

Response Syntax

```
HTTP/1.1 200
Content-type: application/json

{
  "paymentSession": {
    "availableLimits": {
      "availableSpendAmount": {
        "currency": "string",
        "value": "string"
      },
      "updatedAt": "string"
    },
    "createdAt": "string",
    "expiryTimeInMinutes": number,
    "limits": {
      "maxSpendAmount": {
        "currency": "string",
        "value": "string"
      }
    },
    "paymentManagerArn": "string",
    "paymentSessionId": "string",
```

```
"updatedAt": "string",  
"userId": "string"  
}  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

paymentSession

The payment session details.

Type: [PaymentSession](#) object

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

InternalServerError

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetRecommendation

Retrieves detailed information about a recommendation, including its configuration, status, and results.

Request Syntax

```
GET /recommendations/recommendationId HTTP/1.1
```

URI Request Parameters

The request uses the following URI parameters.

recommendationId

The unique identifier of the recommendation to retrieve.

Pattern: `[0-9a-zA-Z_-]{1,48}-[0-9A-Z]{10}`

Required: Yes

Request Body

The request does not have a request body.

Response Syntax

```
HTTP/1.1 200
```

```
Content-type: application/json
```

```
{
  "createdAt": "string",
  "description": "string",
  "kmsKeyArn": "string",
  "name": "string",
  "recommendationArn": "string",
  "recommendationConfig": { ... },
  "recommendationId": "string",
  "recommendationResult": { ... },
  "status": "string",
```

```
"type": "string",  
"updatedAt": "string"  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

createdAt

The timestamp when the recommendation was created.

Type: Timestamp

description

The description of the recommendation.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 4096.

kmsKeyArn

The ARN of the AWS KMS key used to encrypt recommendation data.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: `arn:aws(|-cn|-us-gov):kms:[a-zA-Z0-9-]*:[0-9]{12}:key/[a-zA-Z0-9-]{36}`

name

The name of the recommendation.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 100.

Pattern: `[a-zA-Z][a-zA-Z0-9_-]{0,47}`

recommendationArn

The Amazon Resource Name (ARN) of the recommendation.

Type: String

Pattern: `arn:aws[a-zA-Z-]*:bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:recommendation/[0-9a-zA-Z_-]{1,48}-[0-9A-Z]{10}`

recommendationConfig

The configuration for the recommendation.

Type: [RecommendationConfig](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

recommendationId

The unique identifier of the recommendation.

Type: String

Pattern: `[0-9a-zA-Z_-]{1,48}-[0-9A-Z]{10}`

recommendationResult

The result of the recommendation, containing the optimized system prompt or tool descriptions. Only present when the recommendation status is COMPLETED.

Type: [RecommendationResult](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

status

The current status of the recommendation.

Type: String

Valid Values: PENDING | IN_PROGRESS | COMPLETED | FAILED | DELETING

type

The type of recommendation.

Type: String

Valid Values: SYSTEM_PROMPT_RECOMMENDATION | TOOL_DESCRIPTION_RECOMMENDATION

updatedAt

The timestamp when the recommendation was last updated.

Type: Timestamp

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

InternalServerError

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetResourceApiKey

Retrieves the API key associated with an API key credential provider.

Request Syntax

```
POST /identities/api-key HTTP/1.1
Content-type: application/json

{
  "resourceCredentialProviderName": "string",
  "workloadIdentityToken": "string"
}
```

URI Request Parameters

The request does not use any URI parameters.

Request Body

The request accepts the following data in JSON format.

resourceCredentialProviderName

The credential provider name for the resource from which you are retrieving the API key.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 128.

Pattern: [a-zA-Z0-9\-_]+

Required: Yes

workloadIdentityToken

The identity token of the workload from which you want to retrieve the API key.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 131072.

Required: Yes

Response Syntax

```
HTTP/1.1 200
Content-type: application/json

{
  "apiKey": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

apiKey

The API key associated with the resource requested.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 65536.

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

InternalServerError

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

UnauthorizedException

This exception is thrown when the JWT bearer token is invalid or not found for OAuth bearer token based access

HTTP Status Code: 401

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)

- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetResourceOauth2Token

Returns the OAuth 2.0 token of the provided resource.

Request Syntax

```
POST /identities/oauth2/token HTTP/1.1
Content-type: application/json

{
  "audiences": [ "string" ],
  "customParameters": {
    "string" : "string"
  },
  "customState": "string",
  "forceAuthentication": boolean,
  "oauth2Flow": "string",
  "resourceCredentialProviderName": "string",
  "resourceOauth2ReturnUrl": "string",
  "resources": [ "string" ],
  "scopes": [ "string" ],
  "sessionUri": "string",
  "workloadIdentityToken": "string"
}
```

URI Request Parameters

The request does not use any URI parameters.

Request Body

The request accepts the following data in JSON format.

audiences

The audiences to include in the token request. These are used to specify the intended recipients of the OAuth2 token.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 2048.

Required: No

customParameters

A map of custom parameters to include in the authorization request to the resource credential provider. These parameters are in addition to the standard OAuth 2.0 flow parameters, and will not override them.

Type: String to string map

Key Length Constraints: Minimum length of 1. Maximum length of 256.

Key Pattern: [a-zA-Z0-9\-_\.\]+

Value Length Constraints: Minimum length of 1. Maximum length of 2048.

Required: No

customState

An opaque string that will be sent back to the callback URL provided in resourceOauth2ReturnUrl. This state should be used to protect the callback URL of your application against CSRF attacks by ensuring the response corresponds to the original request.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 4096.

Required: No

forceAuthentication

Indicates whether to always initiate a new three-legged OAuth (3LO) flow, regardless of any existing session.

Type: Boolean

Required: No

oauth2Flow

The type of flow to be performed.

Type: String

Valid Values: USER_FEDERATION | M2M | ON_BEHALF_OF_TOKEN_EXCHANGE

Required: Yes

resourceCredentialProviderName

The name of the resource's credential provider.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 128.

Pattern: [a-zA-Z0-9\-_]+

Required: Yes

resourceOauth2ReturnUrl

The callback URL to redirect to after the OAuth 2.0 token retrieval is complete. This URL must be one of the provided URLs configured for the workload identity.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: \w+ : (\/?\/?)[^\s]+

Required: No

resources

The resources to include in the token request. These are used to specify the target resources for which the OAuth2 token is being requested.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 2048.

Required: No

scopes

The OAuth scopes being requested.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 128.

Required: Yes

sessionUri

Unique identifier for the user's authentication session for retrieving OAuth2 tokens. This ID tracks the authorization flow state across multiple requests and responses during the OAuth2 authentication process.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1024.

Pattern: `urn:iETF:params:oauth:request_uri:[a-zA-Z0-9-._~]+`

Required: No

workloadIdentityToken

The identity token of the workload from which you want to retrieve the OAuth2 token.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 131072.

Required: Yes

Response Syntax

```
HTTP/1.1 200
Content-type: application/json

{
  "accessToken": "string",
  "authorizationUrl": "string",
  "sessionStatus": "string",
  "sessionUri": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

accessToken

The OAuth 2.0 access token to use.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 131072.

authorizationUrl

The URL to initiate the authorization process, provided when the access token requires user authorization.

Type: String

Length Constraints: Minimum length of 1.

sessionStatus

Status indicating whether the user's authorization session is in progress or has failed. This helps determine the next steps in the OAuth2 authentication flow.

Type: String

Valid Values: IN_PROGRESS | FAILED

sessionUri

Unique identifier for the user's authorization session for retrieving OAuth2 tokens. This matches the sessionId from the request and can be used to track the session state.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1024.

Pattern: `urn:ietf:params:oauth:request_uri:[a-zA-Z0-9-._~]+`

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

InternalServerError

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

UnauthorizedException

This exception is thrown when the JWT bearer token is invalid or not found for OAuth bearer token based access

HTTP Status Code: 401

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetResourcePaymentToken

Generates authentication tokens for payment providers that use vendor-specific authentication mechanisms.

Request Syntax

```
POST /identities/payment/token HTTP/1.1
Content-type: application/json

{
  "paymentTokenRequest": { ... },
  "resourceCredentialProviderName": "string",
  "workloadIdentityToken": "string"
}
```

URI Request Parameters

The request does not use any URI parameters.

Request Body

The request accepts the following data in JSON format.

paymentTokenRequest

Vendor-specific token request input. Contains all request parameters in a type-safe, vendor-specific structure.

Type: [PaymentTokenRequestInput](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: Yes

resourceCredentialProviderName

Name of the payment credential provider to use.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 128.

Pattern: `[a-zA-Z0-9\-_]+`

Required: Yes

[workloadIdentityToken](#)

Workload access token for authorization.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 131072.

Required: Yes

Response Syntax

```
HTTP/1.1 200
Content-type: application/json

{
  "paymentTokenResponse": { ... }
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

[paymentTokenResponse](#)

Vendor-specific token response output. Contains all response data in a type-safe, vendor-specific structure.

Type: [PaymentTokenResponseOutput](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

InternalServerError

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

UnauthorizedException

This exception is thrown when the JWT bearer token is invalid or not found for OAuth bearer token based access

HTTP Status Code: 401

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetWorkloadAccessToken

Obtains a workload access token for agentic workloads not acting on behalf of a user.

Request Syntax

```
POST /identities/GetWorkloadAccessToken HTTP/1.1
Content-type: application/json

{
  "workloadName": "string"
}
```

URI Request Parameters

The request does not use any URI parameters.

Request Body

The request accepts the following data in JSON format.

workloadName

The unique identifier for the registered workload.

Type: String

Length Constraints: Minimum length of 3. Maximum length of 255.

Pattern: [A-Za-z0-9_.-]+

Required: Yes

Response Syntax

```
HTTP/1.1 200
Content-type: application/json

{
  "workloadAccessToken": "string"
}
```

```
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

workloadAccessToken

An opaque token representing the identity of both the workload and the user.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 131072.

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

InternalServerError

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

UnauthorizedException

This exception is thrown when the JWT bearer token is invalid or not found for OAuth bearer token based access

HTTP Status Code: 401

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetWorkloadAccessTokenForJWT

Obtains a workload access token for agentic workloads acting on behalf of a user, using a JWT token.

Request Syntax

```
POST /identities/GetWorkloadAccessTokenForJWT HTTP/1.1
```

```
Content-type: application/json
```

```
{
  "userToken": "string",
  "workloadName": "string"
}
```

URI Request Parameters

The request does not use any URI parameters.

Request Body

The request accepts the following data in JSON format.

userToken

The OAuth 2.0 token issued by the user's identity provider.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 131072.

Pattern: `[A-Za-z0-9-_=]+.[A-Za-z0-9-_=]+.[A-Za-z0-9-_=]+`

Required: Yes

workloadName

The unique identifier for the registered workload.

Type: String

Length Constraints: Minimum length of 3. Maximum length of 255.

Pattern: [A-Za-z0-9_.-]+

Required: Yes

Response Syntax

```
HTTP/1.1 200
Content-type: application/json

{
  "workloadAccessToken": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

workloadAccessToken

An opaque token representing the identity of both the workload and the user.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 131072.

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

InternalServerErrorException

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

UnauthorizedException

This exception is thrown when the JWT bearer token is invalid or not found for OAuth bearer token based access

HTTP Status Code: 401

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)

- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetWorkloadAccessTokenForUserId

Obtains a workload access token for agentic workloads acting on behalf of a user, using the user's ID.

Request Syntax

```
POST /identities/GetWorkloadAccessTokenForUserId HTTP/1.1
Content-type: application/json

{
  "userId": "string",
  "workloadName": "string"
}
```

URI Request Parameters

The request does not use any URI parameters.

Request Body

The request accepts the following data in JSON format.

userId

The ID of the user for whom you are retrieving the access token.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 128.

Required: Yes

workloadName

The name of the workload from which you want to retrieve the access token.

Type: String

Length Constraints: Minimum length of 3. Maximum length of 255.

Pattern: [A-Za-z0-9_.-]+

Required: Yes

Response Syntax

```
HTTP/1.1 200
Content-type: application/json

{
  "workloadAccessToken": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

[workloadAccessToken](#)

The access token for the specified workload.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 131072.

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

InternalServerError

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

UnauthorizedException

This exception is thrown when the JWT bearer token is invalid or not found for OAuth bearer token based access

HTTP Status Code: 401

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)

- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

InvokeAgentRuntime

Sends a request to an agent or tool hosted in an Amazon Bedrock AgentCore Runtime and receives responses in real-time.

To invoke an agent, you can specify either the AgentCore Runtime ARN or the agent ID with an account ID, and provide a payload containing your request. When you use the agent ID instead of the full ARN, you don't need to URL-encode the identifier. You can optionally specify a qualifier to target a specific endpoint of the agent.

This operation supports streaming responses, allowing you to receive partial responses as they become available. We recommend using pagination to ensure that the operation returns quickly and successfully when processing large responses.

For example code, see [Invoke an AgentCore Runtime agent](#).

If you're integrating your agent with OAuth, you can't use the AWS SDK to call `InvokeAgentRuntime`. Instead, make a HTTPS request to `InvokeAgentRuntime`. For an example, see [Authenticate and authorize with Inbound Auth and Outbound Auth](#).

To use this operation, you must have the `bedrock-agentcore:InvokeAgentRuntime` permission. If you are making a call to `InvokeAgentRuntime` on behalf of a user ID with the `X-Amzn-Bedrock-AgentCore-Runtime-User-Id` header, You require permissions to both actions (`bedrock-agentcore:InvokeAgentRuntime` and `bedrock-agentcore:InvokeAgentRuntimeForUser`).

Request Syntax

```
POST /runtimes/agentRuntimeArn/invocations?accountId=accountId&qualifier=qualifier
HTTP/1.1
Content-Type: contentType
Accept: accept
Mcp-Session-Id: mcpSessionId
X-Amzn-Bedrock-AgentCore-Runtime-Session-Id: runtimeSessionId
Mcp-Protocol-Version: mcpProtocolVersion
X-Amzn-Bedrock-AgentCore-Runtime-User-Id: runtimeUserId
X-Amzn-Trace-Id: traceId
traceparent: traceParent
tracestate: traceState
baggage: baggage
```

payload

URI Request Parameters

The request uses the following URI parameters.

accept

The desired MIME type for the response from the agent runtime. This tells the agent runtime what format to use for the response data. Common values include `application/json` for JSON data.

Length Constraints: Minimum length of 1. Maximum length of 256.

accountId

The identifier of the AWS account for the agent runtime resource. This parameter is required when you specify an agent ID instead of the full ARN for `agentRuntimeArn`.

Pattern: `[0-9]{12}`

agentRuntimeArn

The identifier of the agent runtime to invoke. You can specify either the full AWS Resource Name (ARN) or the agent ID. If you use the agent ID, you must also provide the `accountId` query parameter.

Required: Yes

baggage

Additional context information for distributed tracing.

Length Constraints: Minimum length of 0. Maximum length of 8192.

contentType

The MIME type of the input data in the payload. This tells the agent runtime how to interpret the payload data. Common values include `application/json` for JSON data.

Length Constraints: Minimum length of 1. Maximum length of 256.

mcpProtocolVersion

The version of the MCP protocol being used.

Length Constraints: Minimum length of 1. Maximum length of 1024.

mcpSessionId

The identifier of the MCP session.

Length Constraints: Minimum length of 1. Maximum length of 1024.

qualifier

The qualifier to use for the agent runtime. This is an endpoint name that points to a specific version. If not specified, Amazon Bedrock AgentCore uses the default endpoint of the agent runtime.

runtimeSessionId

The identifier of the runtime session.

Length Constraints: Minimum length of 33. Maximum length of 256.

runtimeUserId

The identifier of the runtime user.

Length Constraints: Minimum length of 1. Maximum length of 1024.

traceId

The trace identifier for request tracking.

Length Constraints: Minimum length of 0. Maximum length of 128.

traceParent

The parent trace information for distributed tracing.

Length Constraints: Minimum length of 0. Maximum length of 128.

traceState

The trace state information for distributed tracing.

Length Constraints: Minimum length of 0. Maximum length of 512.

Request Body

The request accepts the following binary data.

payload

The input data to send to the agent runtime. The format of this data depends on the specific agent configuration and must match the specified content type. For most agents, this is a JSON object containing the user's request.

Length Constraints: Minimum length of 0. Maximum length of 100000000.

Required: Yes

Response Syntax

```
HTTP/1.1 statusCode
X-Amzn-Bedrock-AgentCore-Runtime-Session-Id: runtimeSessionId
Mcp-Session-Id: mcpSessionId
Mcp-Protocol-Version: mcpProtocolVersion
X-Amzn-Trace-Id: traceId
traceparent: traceParent
tracestate: traceState
baggage: baggage
Content-Type: contentType

response
```

Response Elements

If the action is successful, the service sends back the following HTTP response.

statusCode

The HTTP status code of the response. A status code of 200 indicates a successful operation. Other status codes indicate various error conditions.

The response returns the following HTTP headers.

baggage

Additional context information for distributed tracing.

contentType

The MIME type of the response data. This indicates how to interpret the response data. Common values include application/json for JSON data.

mcpProtocolVersion

The version of the MCP protocol being used.

mcpSessionId

The identifier of the MCP session.

Length Constraints: Minimum length of 1. Maximum length of 100.

Pattern: [a-zA-Z0-9][a-zA-Z0-9-_*]

runtimeSessionId

The identifier of the runtime session.

Length Constraints: Minimum length of 1. Maximum length of 100.

Pattern: [a-zA-Z0-9][a-zA-Z0-9-_*]

traceId

The trace identifier for request tracking.

traceParent

The parent trace information for distributed tracing.

traceState

The trace state information for distributed tracing.

The response returns the following as the HTTP body.

response

The response data from the agent runtime. The format of this data depends on the specific agent configuration and the requested accept type. For most agents, this is a JSON object containing the agent's response to the user's request.

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

InternalServerError

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

RetryableConflictException

The exception that occurs when there is a retryable conflict performing an operation. This is a temporary condition that may resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 409

RuntimeClientError

The exception that occurs when there is an error in the runtime client. This can happen due to network issues, invalid configuration, or other client-side problems. Check the error message for specific details about the error.

HTTP Status Code: 424

ServiceQuotaExceededException

The exception that occurs when the request would cause a service quota to be exceeded. Review your service quotas and either reduce your request rate or request a quota increase.

HTTP Status Code: 402

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

InvokeAgentRuntimeCommand

Executes a command in a runtime session container and streams the output back to the caller. This operation allows you to run shell commands within the agent runtime environment and receive real-time streaming responses including standard output and standard error.

To invoke a command, you must specify the agent runtime ARN and a runtime session ID. The command execution supports streaming responses, allowing you to receive output as it becomes available through `contentStart`, `contentDelta`, and `contentStop` events.

To use this operation, you must have the `bedrock-agentcore:InvokeAgentRuntimeCommand` permission.

Request Syntax

```
POST /runtimes/agentRuntimeArn/commands?accountId=accountId&qualifier=qualifier
HTTP/1.1
Content-Type: contentType
Accept: accept
X-Amzn-Bedrock-AgentCore-Runtime-Session-Id: runtimeSessionId
X-Amzn-Trace-Id: traceId
traceparent: traceParent
tracestate: traceState
baggage: baggage
Content-type: application/json

{
  "command": "string",
  "timeout": number
}
```

URI Request Parameters

The request uses the following URI parameters.

accept

The desired MIME type for the response from the agent runtime command. This tells the agent runtime what format to use for the response data. Common values include `application/json` for JSON data.

Length Constraints: Minimum length of 1. Maximum length of 256.

accountId

The identifier of the AWS account for the agent runtime resource. This parameter is required when you specify an agent ID instead of the full ARN for `agentRuntimeArn`.

Pattern: `[0-9]{12}`

agentRuntimeArn

The Amazon Resource Name (ARN) of the agent runtime on which to execute the command. This identifies the specific agent runtime environment where the command will run.

Required: Yes

baggage

Additional context information for distributed tracing.

Length Constraints: Minimum length of 0. Maximum length of 8192.

contentType

The MIME type of the input data in the request payload. This tells the agent runtime how to interpret the payload data. Common values include `application/json` for JSON data.

Length Constraints: Minimum length of 1. Maximum length of 256.

qualifier

The qualifier to use for the agent runtime. This is an endpoint name that points to a specific version. If not specified, Amazon Bedrock AgentCore uses the default endpoint of the agent runtime.

runtimeSessionId

The unique identifier of the runtime session in which to execute the command. This session ID is used to maintain state and context across multiple command invocations.

Length Constraints: Minimum length of 33. Maximum length of 256.

traceId

The trace identifier for request tracking.

Length Constraints: Minimum length of 0. Maximum length of 1024.

traceParent

The parent trace information for distributed tracing.

Length Constraints: Minimum length of 0. Maximum length of 1024.

traceState

The trace state information for distributed tracing.

Length Constraints: Minimum length of 0. Maximum length of 512.

Request Body

The request accepts the following data in JSON format.

command

The shell command to execute on the agent runtime. This command is executed in the runtime environment and its output is streamed back to the caller.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 65536.

Required: Yes

timeout

The maximum duration in seconds to wait for the command to complete. If the command execution exceeds this timeout, it will be terminated. Default is 300 seconds. Minimum is 1 second. Maximum is 3600 seconds.

Type: Integer

Required: No

Response Syntax

```
HTTP/1.1 statusCode  
X-Amzn-Bedrock-AgentCore-Runtime-Session-Id: runtimeSessionId  
X-Amzn-Trace-Id: traceId
```

```
traceparent: traceParent
tracestate: traceState
baggage: baggage
Content-Type: contentType
Content-type: application/json
```

```
{
  "accessDeniedException": {
  },
  "chunk": {
    "contentDelta": {
      "stderr": "string",
      "stdout": "string"
    },
    "contentStart": {
    },
    "contentStop": {
      "exitCode": number,
      "status": "string"
    }
  },
  "internalServerError": {
  },
  "resourceNotFoundException": {
  },
  "runtimeClientError": {
  },
  "serviceQuotaExceededException": {
  },
  "throttlingException": {
  },
  "validationException": {
  }
}
```

Response Elements

If the action is successful, the service sends back the following HTTP response.

[statusCode](#)

The HTTP status code of the response. A status code of 200 indicates a successful operation. Other status codes indicate various error conditions.

The response returns the following HTTP headers.

baggage

Additional context information for distributed tracing.

contentType

The MIME type of the response data. This indicates how to interpret the response data. Common values include application/json for JSON data.

runtimeSessionId

The unique identifier of the runtime session in which the command was executed.

Length Constraints: Minimum length of 1. Maximum length of 100.

Pattern: [a-zA-Z0-9][a-zA-Z0-9-_*]

traceId

The trace identifier for request tracking.

traceParent

The parent trace information for distributed tracing.

traceState

The trace state information for distributed tracing.

The following data is returned in JSON format by the service.

accessDeniedException

Exception events for error streaming.

Type: Exception

HTTP Status Code: 403

chunk

A response chunk containing command execution events such as content start, content delta, or content stop events.

Type: [ResponseChunk](#) object

internalServerErrorException

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

Type: Exception

HTTP Status Code: 500

resourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

Type: Exception

HTTP Status Code: 404

runtimeClientError

The exception that occurs when there is an error in the runtime client. This can happen due to network issues, invalid configuration, or other client-side problems. Check the error message for specific details about the error.

Type: Exception

HTTP Status Code: 424

serviceQuotaExceededException

The exception that occurs when the request would cause a service quota to be exceeded. Review your service quotas and either reduce your request rate or request a quota increase.

Type: Exception

HTTP Status Code: 402

throttlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

Type: Exception

HTTP Status Code: 429

validationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

Type: Exception

HTTP Status Code: 400

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

InternalServerError

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

RuntimeClientError

The exception that occurs when there is an error in the runtime client. This can happen due to network issues, invalid configuration, or other client-side problems. Check the error message for specific details about the error.

HTTP Status Code: 424

ServiceQuotaExceededException

The exception that occurs when the request would cause a service quota to be exceeded. Review your service quotas and either reduce your request rate or request a quota increase.

HTTP Status Code: 402

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)

- [AWS SDK for Ruby V3](#)

InvokeBrowser

Invokes an operating system-level action on a browser session in Amazon Bedrock AgentCore. This operation provides direct OS-level control over browser sessions, enabling mouse actions, keyboard input, and screenshots that the WebSocket-based Chrome DevTools Protocol (CDP) cannot handle — such as interacting with print dialogs, context menus, and JavaScript alerts.

You send a request with exactly one action in the `BrowserAction` union, and receive a corresponding result in the `BrowserActionResult` union.

The following operations are related to `InvokeBrowser`:

- [StartBrowserSession](#)
- [GetBrowserSession](#)
- [StopBrowserSession](#)

Request Syntax

```
POST /browsers/browserIdentifier/sessions/invoke HTTP/1.1
x-amzn-browser-session-id: sessionId
Content-type: application/json

{
  "action": { ... }
}
```

URI Request Parameters

The request uses the following URI parameters.

[browserIdentifier](#)

The unique identifier of the browser associated with the session. This must match the identifier used when creating the session with `StartBrowserSession`.

Required: Yes

[sessionId](#)

The unique identifier of the browser session on which to perform the action. This must be an active session created with `StartBrowserSession`.

Pattern: `[0-9a-zA-Z]{1,40}`

Required: Yes

Request Body

The request accepts the following data in JSON format.

action

The browser action to perform. Exactly one member of the `BrowserAction` union must be set per request.

Type: [BrowserAction](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: Yes

Response Syntax

```
HTTP/1.1 200
x-amzn-browser-session-id: sessionId
Content-type: application/json

{
  "result": { ... }
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The response returns the following HTTP headers.

sessionId

The unique identifier of the browser session on which the action was performed.

Pattern: `[0-9a-zA-Z]{1,40}`

The following data is returned in JSON format by the service.

result

The result of the browser action. The member set in the result corresponds to the action that was performed.

Type: [BrowserActionResult](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

InternalServerError

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ServiceQuotaExceededException

The exception that occurs when the request would cause a service quota to be exceeded. Review your service quotas and either reduce your request rate or request a quota increase.

HTTP Status Code: 402

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

InvokeCodeInterpreter

Executes code within an active code interpreter session in Amazon Bedrock AgentCore. This operation processes the provided code, runs it in a secure environment, and returns the execution results including output, errors, and generated visualizations.

To execute code, you must specify the code interpreter identifier, session ID, and the code to run in the arguments parameter. The operation returns a stream containing the execution results, which can include text output, error messages, and data visualizations.

This operation is subject to request rate limiting based on your account's service quotas.

The following operations are related to InvokeCodeInterpreter:

- [StartCodeInterpreterSession](#)
- [GetCodeInterpreterSession](#)

Request Syntax

```
POST /code-interpreters/codeInterpreterIdentifier/tools/invoke HTTP/1.1
```

```
x-amzn-code-interpreter-session-id: sessionId
```

```
X-Amzn-Trace-Id: traceId
```

```
traceparent: traceParent
```

```
Content-type: application/json
```

```
{
  "arguments": {
    "clearContext": boolean,
    "code": "string",
    "command": "string",
    "content": [
      {
        "blob": blob,
        "path": "string",
        "text": "string"
      }
    ],
    "directoryPath": "string",
    "language": "string",
    "path": "string",
```

```
  "paths": [ "string" ],
  "runtime": "string",
  "taskId": "string"
},
"name": "string"
}
```

URI Request Parameters

The request uses the following URI parameters.

[codeInterpreterIdentifier](#)

The unique identifier of the code interpreter associated with the session. This must match the identifier used when creating the session with `StartCodeInterpreterSession`.

Required: Yes

[sessionId](#)

The unique identifier of the code interpreter session to use. This must be an active session created with `StartCodeInterpreterSession`. If the session has expired or been stopped, the request will fail.

Pattern: `[0-9a-zA-Z]{1,40}`

[traceId](#)

The trace identifier for request tracking.

Length Constraints: Minimum length of 0. Maximum length of 1024.

[traceParent](#)

The parent trace information for distributed tracing.

Length Constraints: Minimum length of 0. Maximum length of 1024.

Request Body

The request accepts the following data in JSON format.

arguments

The arguments for the code interpreter. This includes the code to execute and any additional parameters such as the programming language, whether to clear the execution context, and other execution options. The structure of this parameter depends on the specific code interpreter being used.

Type: [ToolArguments](#) object

Required: No

name

The name of the code interpreter to invoke.

Type: String

Valid Values: `executeCode` | `executeCommand` | `readFiles` | `listFiles` | `removeFiles` | `writeFiles` | `startCommandExecution` | `getTask` | `stopTask`

Required: Yes

Response Syntax

```
HTTP/1.1 200
x-amzn-code-interpreter-session-id: sessionId
Content-type: application/json
```

```
{
  "accessDeniedException": {
  },
  "conflictException": {
  },
  "internalServerErrorException": {
  },
  "resourceNotFoundException": {
  },
  "result": {
    "content": [
      {
        "data": blob,
        "description": "string",
        "mimeType": "string",
```

```
    "name": "string",
    "resource": {
      "blob": blob,
      "mimeType": "string",
      "text": "string",
      "type": "string",
      "uri": "string"
    },
    "size": number,
    "text": "string",
    "type": "string",
    "uri": "string"
  }
],
"isError": boolean,
"structuredContent": {
  "executionTime": number,
  "exitCode": number,
  "stderr": "string",
  "stdout": "string",
  "taskId": "string",
  "taskStatus": "string"
}
},
"serviceQuotaExceededException": {
},
"throttlingException": {
},
"validationException": {
}
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The response returns the following HTTP headers.

sessionId

The identifier of the code interpreter session.

Pattern: `[0-9a-zA-Z]{1,40}`

The following data is returned in JSON format by the service.

[accessDeniedException](#)

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

Type: Exception

HTTP Status Code: 403

[conflictException](#)

The exception that occurs when the request conflicts with the current state of the resource. This can happen when trying to modify a resource that is currently being modified by another request, or when trying to create a resource that already exists.

Type: Exception

HTTP Status Code: 409

[internalServerErrorException](#)

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

Type: Exception

HTTP Status Code: 500

[resourceNotFoundException](#)

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

Type: Exception

HTTP Status Code: 404

[result](#)

The output produced by executing code in a code interpreter session in Amazon Bedrock AgentCore. This structure contains the results of code execution, including textual output,

structured data, and error information. Agents use these results to generate responses that incorporate computation, data analysis, and visualization.

Type: [CodeInterpreterResult](#) object

[serviceQuotaExceededException](#)

The exception that occurs when the request would cause a service quota to be exceeded. Review your service quotas and either reduce your request rate or request a quota increase.

Type: Exception

HTTP Status Code: 402

[throttlingException](#)

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

Type: Exception

HTTP Status Code: 429

[validationException](#)

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

Type: Exception

HTTP Status Code: 400

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

ConflictException

The exception that occurs when the request conflicts with the current state of the resource. This can happen when trying to modify a resource that is currently being modified by another request, or when trying to create a resource that already exists.

HTTP Status Code: 409

InternalServerErrorException

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ServiceQuotaExceededException

The exception that occurs when the request would cause a service quota to be exceeded. Review your service quotas and either reduce your request rate or request a quota increase.

HTTP Status Code: 402

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

InvokeHarness

Operation to invoke a Harness.

Request Syntax

```
POST /harnesses/invoke?harnessArn=harnessArn&qualifier=qualifier HTTP/1.1
```

```
X-Amzn-Bedrock-AgentCore-Runtime-Session-Id: runtimeSessionId
```

```
X-Amzn-Bedrock-AgentCore-Runtime-User-Id: runtimeUserId
```

```
Content-type: application/json
```

```
{
  "actorId": "string",
  "allowedTools": [ "string" ],
  "maxIterations": number,
  "maxTokens": number,
  "messages": [
    {
      "content": [
        { ... }
      ],
      "role": "string"
    }
  ],
  "model": { ... },
  "skills": [
    { ... }
  ],
  "systemPrompt": [
    { ... }
  ],
  "timeoutSeconds": number,
  "tools": [
    {
      "config": { ... },
      "name": "string",
      "type": "string"
    }
  ]
}
```

URI Request Parameters

The request uses the following URI parameters.

[harnessArn](#)

The ARN of the harness to invoke.

Pattern: `arn:([^:]+)?:bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:harness/[a-zA-Z][a-zA-Z0-9_]{0,39}-[a-zA-Z0-9]{10}`

Required: Yes

[qualifier](#)

The endpoint name to invoke. If omitted, the DEFAULT endpoint is used.

Pattern: `[a-zA-Z][a-zA-Z0-9_]{0,47}`

[runtimeSessionId](#)

The session ID for the invocation. Use the same session ID across requests to continue a conversation.

Length Constraints: Minimum length of 33. Maximum length of 100.

Pattern: `[a-zA-Z0-9][a-zA-Z0-9-_*]`

Required: Yes

[runtimeUserId](#)

An identifier for the end user making the request. This value is passed through to the runtime container.

Request Body

The request accepts the following data in JSON format.

[actorId](#)

The actor ID for memory operations. Overrides the actor ID configured on the harness.

Type: String

Required: No

allowedTools

The tools that the agent is allowed to use for this invocation. If specified, overrides the harness default.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: (`* |@?[^/]+([^/]+)?`)

Required: No

maxIterations

The maximum number of iterations the agent loop can execute. If specified, overrides the harness default.

Type: Integer

Required: No

maxTokens

The maximum number of tokens the agent can generate per iteration. If specified, overrides the harness default.

Type: Integer

Required: No

messages

The messages to send to the agent.

Type: Array of [HarnessMessage](#) objects

Required: Yes

model

The model configuration to use for this invocation. If specified, overrides the harness default.

Type: [HarnessModelConfiguration](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: No

skills

The skills available to the agent for this invocation. If specified, overrides the harness default.

Type: Array of [HarnessSkill](#) objects

Required: No

systemPrompt

The system prompt to use for this invocation. If specified, overrides the harness default.

Type: Array of [HarnessSystemContentBlock](#) objects

Required: No

timeoutSeconds

The maximum duration in seconds for the agent loop execution. If specified, overrides the harness default.

Type: Integer

Required: No

tools

The tools available to the agent for this invocation. If specified, overrides the harness default.

Type: Array of [HarnessTool](#) objects

Required: No

Response Syntax

```
HTTP/1.1 200
Content-type: application/json

{
  "contentBlockDelta": {
    "contentBlockIndex": number,
    "delta": { ... }
  }
}
```

```
},
  "contentBlockStart": {
    "contentBlockIndex": number,
    "start": { ... }
  },
  "contentBlockStop": {
    "contentBlockIndex": number
  },
  "internalServerError": {
  },
  "messageStart": {
    "role": "string"
  },
  "messageStop": {
    "stopReason": "string"
  },
  "metadata": {
    "metrics": {
      "latencyMs": number
    },
    "usage": {
      "cacheReadInputTokens": number,
      "cacheWriteInputTokens": number,
      "inputTokens": number,
      "outputTokens": number,
      "totalTokens": number
    }
  },
  "runtimeClientError": {
  },
  "validationException": {
  }
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

contentBlockDelta

A delta update to the current content block.

Type: [HarnessContentBlockDeltaEvent](#) object

[contentBlockStart](#)

Indicates the start of a new content block.

Type: [HarnessContentBlockStartEvent](#) object

[contentBlockStop](#)

Indicates the end of the current content block.

Type: [HarnessContentBlockStopEvent](#) object

[internalServerError](#)

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

Type: Exception

HTTP Status Code: 500

[messageStart](#)

Indicates the start of a new message from the agent.

Type: [HarnessMessageStartEvent](#) object

[messageStop](#)

Indicates the end of the current message.

Type: [HarnessMessageStopEvent](#) object

[metadata](#)

Token usage and latency metrics for the invocation.

Type: [HarnessMetadataEvent](#) object

[runtimeClientError](#)

An error returned by the runtime container during agent execution.

Type: Exception

HTTP Status Code: 424

validationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

Type: Exception

HTTP Status Code: 400

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

InternalServerError

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

RuntimeClientError

The exception that occurs when there is an error in the runtime client. This can happen due to network issues, invalid configuration, or other client-side problems. Check the error message for specific details about the error.

HTTP Status Code: 424

ServiceQuotaExceededException

The exception that occurs when the request would cause a service quota to be exceeded. Review your service quotas and either reduce your request rate or request a quota increase.

HTTP Status Code: 402

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)

- [AWS SDK for Ruby V3](#)

ListABTests

Lists all A/B tests in the account.

Request Syntax

```
GET /ab-tests?maxResults=maxResults&nextToken=nextToken HTTP/1.1
```

URI Request Parameters

The request uses the following URI parameters.

[maxResults](#)

The maximum number of results to return in the response. If the total number of results is greater than this value, use the token returned in the response in the `nextToken` field when making another request to return the next batch of results.

Valid Range: Minimum value of 1. Maximum value of 100.

[nextToken](#)

If the total number of results is greater than the `maxResults` value provided in the request, enter the token returned in the `nextToken` field in the response in this field to return the next batch of results.

Request Body

The request does not have a request body.

Response Syntax

```
HTTP/1.1 200
Content-type: application/json

{
  "abTests": [
    {
      "abTestArn": "string",
```

```
    "abTestId": "string",
    "createdAt": number,
    "description": "string",
    "executionStatus": "string",
    "gatewayArn": "string",
    "name": "string",
    "status": "string",
    "updatedAt": number
  }
],
"nextToken": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

abTests

The list of A/B test summaries.

Type: Array of [ABTestSummary](#) objects

nextToken

If the total number of results is greater than the `maxResults` value provided in the request, use this token when making another request in the `nextToken` field to return the next batch of results.

Type: String

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

InternalServerErrorException

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

UnauthorizedException

This exception is thrown when the JWT bearer token is invalid or not found for OAuth bearer token based access

HTTP Status Code: 401

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)

- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

ListActors

Lists all actors in an AgentCore Memory resource. We recommend using pagination to ensure that the operation returns quickly and successfully.

To use this operation, you must have the `bedrock-agentcore:ListActors` permission.

Request Syntax

```
POST /memories/memoryId/actors HTTP/1.1
```

```
Content-type: application/json
```

```
{
  "maxResults": number,
  "nextToken": "string"
}
```

URI Request Parameters

The request uses the following URI parameters.

memoryId

The identifier of the AgentCore Memory resource for which to list actors.

Length Constraints: Minimum length of 12.

Pattern: `(arn:(aws|aws-cn|aws-us-gov):bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:memory/)?[a-zA-Z][a-zA-Z0-9-_{0,99}]-[a-zA-Z0-9]{10}`

Required: Yes

Request Body

The request accepts the following data in JSON format.

maxResults

The maximum number of results to return in a single call. The default value is 20.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

[nextToken](#)

The token for the next set of results. Use the value returned in the previous response in the next request to retrieve the next set of results.

Type: String

Required: No

Response Syntax

```
HTTP/1.1 200
Content-type: application/json

{
  "actorSummaries": [
    {
      "actorId": "string"
    }
  ],
  "nextToken": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

[actorSummaries](#)

The list of actor summaries.

Type: Array of [ActorSummary](#) objects

[nextToken](#)

The token to use in a subsequent request to get the next set of results. This value is null when there are no more results to return.

Type: String

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

InvalidInputException

The input fails to satisfy the constraints specified by AgentCore. Check your input values and try again.

HTTP Status Code: 400

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ServiceException

The service encountered an internal error. Try your request again later.

HTTP Status Code: 500

ServiceQuotaExceededException

The exception that occurs when the request would cause a service quota to be exceeded. Review your service quotas and either reduce your request rate or request a quota increase.

HTTP Status Code: 402

ThrottledException

The request was denied due to request throttling. Reduce the frequency of requests and try again.

HTTP Status Code: 429

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

ListBatchEvaluations

Lists all batch evaluations in the account, providing summary information about each evaluation's status and configuration.

Request Syntax

```
GET /evaluations/batch-evaluate?maxResults=maxResults&nextToken=nextToken HTTP/1.1
```

URI Request Parameters

The request uses the following URI parameters.

[maxResults](#)

The maximum number of results to return in the response. If the total number of results is greater than this value, use the token returned in the response in the nextToken field when making another request to return the next batch of results.

Valid Range: Minimum value of 1. Maximum value of 100.

[nextToken](#)

If the total number of results is greater than the maxResults value provided in the request, enter the token returned in the nextToken field in the response in this field to return the next batch of results.

Request Body

The request does not have a request body.

Response Syntax

```
HTTP/1.1 200
Content-type: application/json

{
  "batchEvaluations": [
    {
      "batchEvaluationArn": "string",
```

```
"batchEvaluationId": "string",
"batchEvaluationName": "string",
"createdAt": "string",
"description": "string",
"errorDetails": [ "string" ],
"evaluationResults": {
  "evaluatorSummaries": [
    {
      "evaluatorId": "string",
      "statistics": {
        "averageScore": number
      },
      "totalEvaluated": number,
      "totalFailed": number
    }
  ],
  "numberOfSessionsCompleted": number,
  "numberOfSessionsFailed": number,
  "numberOfSessionsIgnored": number,
  "numberOfSessionsInProgress": number,
  "totalNumberOfSessions": number
},
"evaluators": [
  {
    "evaluatorId": "string"
  }
],
"insights": [
  {
    "insightId": "string"
  }
],
"kmsKeyArn": "string",
"status": "string",
"updatedAt": "string"
}
],
"nextToken": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

[batchEvaluations](#)

The list of batch evaluation summaries.

Type: Array of [BatchEvaluationSummary](#) objects

[nextToken](#)

If the total number of results is greater than the `maxResults` value provided in the request, use this token when making another request in the `nextToken` field to return the next batch of results.

Type: String

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

InternalServerError

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

UnauthorizedException

This exception is thrown when the JWT bearer token is invalid or not found for OAuth bearer token based access

HTTP Status Code: 401

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

ListBrowserSessions

Retrieves a list of browser sessions in Amazon Bedrock AgentCore that match the specified criteria. This operation returns summary information about each session, including identifiers, status, and timestamps.

You can filter the results by browser identifier and session status. The operation supports pagination to handle large result sets efficiently.

We recommend using pagination to ensure that the operation returns quickly and successfully when retrieving large numbers of sessions.

The following operations are related to `ListBrowserSessions`:

- [StartBrowserSession](#)
- [GetBrowserSession](#)

Request Syntax

```
POST /browsers/browserIdentifier/sessions/list HTTP/1.1
Content-type: application/json
```

```
{
  "maxResults": number,
  "nextToken": "string",
  "status": "string"
}
```

URI Request Parameters

The request uses the following URI parameters.

[browserIdentifier](#)

The unique identifier of the browser to list sessions for. If specified, only sessions for this browser are returned. If not specified, sessions for all browsers are returned.

Required: Yes

Request Body

The request accepts the following data in JSON format.

maxResults

The maximum number of results to return in a single call. The default value is 10. Valid values range from 1 to 100. To retrieve the remaining results, make another call with the returned `nextToken` value.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

nextToken

The token for the next set of results. Use the value returned in the previous response in the next request to retrieve the next set of results. If not specified, Amazon Bedrock AgentCore returns the first page of results.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: `\S*`

Required: No

status

The status of the browser sessions to list. Valid values include `ACTIVE`, `STOPPING`, and `STOPPED`. If not specified, sessions with any status are returned.

Type: String

Valid Values: `READY` | `TERMINATED`

Required: No

Response Syntax

```
HTTP/1.1 200
```

```
Content-type: application/json

{
  "items": [
    {
      "browserIdentifier": "string",
      "createdAt": "string",
      "lastUpdatedAt": "string",
      "name": "string",
      "sessionId": "string",
      "status": "string"
    }
  ],
  "nextToken": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

items

The list of browser sessions that match the specified criteria.

Type: Array of [BrowserSessionSummary](#) objects

nextToken

The token to use in a subsequent `ListBrowserSessions` request to get the next set of results.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: `\S*`

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

InternalServerError

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

ListCodeInterpreterSessions

Retrieves a list of code interpreter sessions in Amazon Bedrock AgentCore that match the specified criteria. This operation returns summary information about each session, including identifiers, status, and timestamps.

You can filter the results by code interpreter identifier and session status. The operation supports pagination to handle large result sets efficiently.

We recommend using pagination to ensure that the operation returns quickly and successfully when retrieving large numbers of sessions.

The following operations are related to ListCodeInterpreterSessions:

- [StartCodeInterpreterSession](#)
- [GetCodeInterpreterSession](#)

Request Syntax

```
POST /code-interpreters/codeInterpreterIdentifier/sessions/list HTTP/1.1
Content-type: application/json
```

```
{
  "maxResults": number,
  "nextToken": "string",
  "status": "string"
}
```

URI Request Parameters

The request uses the following URI parameters.

[codeInterpreterIdentifier](#)

The unique identifier of the code interpreter to list sessions for. If specified, only sessions for this code interpreter are returned. If not specified, sessions for all code interpreters are returned.

Required: Yes

Request Body

The request accepts the following data in JSON format.

maxResults

The maximum number of results to return in a single call. The default value is 10. Valid values range from 1 to 100. To retrieve the remaining results, make another call with the returned `nextToken` value.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

nextToken

The token for the next set of results. Use the value returned in the previous response in the next request to retrieve the next set of results. If not specified, Amazon Bedrock AgentCore returns the first page of results.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: `\S*`

Required: No

status

The status of the code interpreter sessions to list. Valid values include `ACTIVE`, `STOPPING`, and `STOPPED`. If not specified, sessions with any status are returned.

Type: String

Valid Values: `READY` | `TERMINATED`

Required: No

Response Syntax

```
HTTP/1.1 200
```

```
Content-type: application/json

{
  "items": [
    {
      "codeInterpreterIdentifier": "string",
      "createdAt": "string",
      "lastUpdatedAt": "string",
      "name": "string",
      "sessionId": "string",
      "status": "string"
    }
  ],
  "nextToken": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

items

The list of code interpreter sessions that match the specified criteria.

Type: Array of [CodeInterpreterSessionSummary](#) objects

nextToken

The token to use in a subsequent `ListCodeInterpreterSessions` request to get the next set of results.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: `\S*`

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

InternalServerError

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

ListEvents

Lists events in an AgentCore Memory resource based on specified criteria. We recommend using pagination to ensure that the operation returns quickly and successfully.

To use this operation, you must have the `bedrock-agentcore:ListEvents` permission.

Request Syntax

```
POST /memories/memoryId/actor/actorId/sessions/sessionId HTTP/1.1
```

```
Content-type: application/json
```

```
{
  "filter": {
    "branch": {
      "includeParentBranches": boolean,
      "name": "string"
    },
    "eventMetadata": [
      {
        "left": { ... },
        "operator": "string",
        "right": { ... }
      }
    ]
  },
  "includePayloads": boolean,
  "maxResults": number,
  "nextToken": "string"
}
```

URI Request Parameters

The request uses the following URI parameters.

actorId

The identifier of the actor for which to list events.

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[a-zA-Z0-9][a-zA-Z0-9-_/]*(?::[a-zA-Z0-9-_/]+)*[a-zA-Z0-9-_/]*`

Required: Yes

memoryId

The identifier of the AgentCore Memory resource for which to list events.

Length Constraints: Minimum length of 12.

Pattern: `(arn:(aws|aws-cn|aws-us-gov):bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:memory/)?[a-zA-Z][a-zA-Z0-9-_{0,99}-[a-zA-Z0-9]{10}`

Required: Yes

sessionId

The identifier of the session for which to list events.

Length Constraints: Minimum length of 1. Maximum length of 100.

Pattern: `[a-zA-Z0-9][a-zA-Z0-9-_*]`

Required: Yes

Request Body

The request accepts the following data in JSON format.

filter

Filter criteria to apply when listing events.

Type: [FilterInput](#) object

Required: No

includePayloads

Specifies whether to include event payloads in the response. Set to true to include payloads, or false to exclude them.

Type: Boolean

Required: No

maxResults

The maximum number of results to return in a single call. The default value is 20.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

nextToken

The token for the next set of results. Use the value returned in the previous response in the next request to retrieve the next set of results.

Type: String

Required: No

Response Syntax

```
HTTP/1.1 200
Content-type: application/json

{
  "events": [
    {
      "actorId": "string",
      "branch": {
        "name": "string",
        "rootEventId": "string"
      },
      "eventId": "string",
      "eventTimestamp": number,
      "memoryId": "string",
      "metadata": {
        "string" : { ... }
      },
      "payload": [
        { ... }
      ],
      "sessionId": "string"
    }
  ]
}
```

```
  ],  
  "nextToken": "string"  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

events

The list of events that match the specified criteria.

Type: Array of [Event](#) objects

nextToken

The token to use in a subsequent request to get the next set of results. This value is null when there are no more results to return.

Type: String

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

InvalidInputException

The input fails to satisfy the constraints specified by AgentCore. Check your input values and try again.

HTTP Status Code: 400

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ServiceException

The service encountered an internal error. Try your request again later.

HTTP Status Code: 500

ServiceQuotaExceededException

The exception that occurs when the request would cause a service quota to be exceeded. Review your service quotas and either reduce your request rate or request a quota increase.

HTTP Status Code: 402

ThrottledException

The request was denied due to request throttling. Reduce the frequency of requests and try again.

HTTP Status Code: 429

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)

- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

ListMemoryExtractionJobs

Lists all long-term memory extraction jobs that are eligible to be started with optional filtering.

To use this operation, you must have the `bedrock-agentcore:ListMemoryExtractionJobs` permission.

Request Syntax

```
POST /memories/memoryId/extractionJobs HTTP/1.1
```

```
Content-type: application/json
```

```
{
  "filter": {
    "actorId": "string",
    "sessionId": "string",
    "status": "string",
    "strategyId": "string"
  },
  "maxResults": number,
  "nextToken": "string"
}
```

URI Request Parameters

The request uses the following URI parameters.

memoryId

The unique identifier of the memory to list extraction jobs for.

Length Constraints: Minimum length of 12.

Pattern: `(arn:(aws|aws-cn|aws-us-gov):bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:memory/)?[a-zA-Z][a-zA-Z0-9-_{0,99}]-[a-zA-Z0-9]{10}`

Required: Yes

Request Body

The request accepts the following data in JSON format.

filter

Filter criteria to apply when listing extraction jobs.

Type: [ExtractionJobFilterInput](#) object

Required: No

maxResults

The maximum number of results to return in a single call. The default value is 20.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 50.

Required: No

nextToken

The token for the next set of results. Use the value returned in the previous response in the next request to retrieve the next set of results.

Type: String

Required: No

Response Syntax

```
HTTP/1.1 200
Content-type: application/json

{
  "jobs": [
    {
      "actorId": "string",
      "failureReason": "string",
      "jobID": "string",
      "messages": { ... },
      "sessionId": "string",
      "status": "string",
      "strategyId": "string"
    }
  ],
}
```

```
"nextToken": "string"  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

jobs

List of extraction job metadata matching the specified criteria.

Type: Array of [ExtractionJobMetadata](#) objects

nextToken

Token to retrieve the next page of results, if available.

Type: String

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ServiceException

The service encountered an internal error. Try your request again later.

HTTP Status Code: 500

ServiceQuotaExceededException

The exception that occurs when the request would cause a service quota to be exceeded. Review your service quotas and either reduce your request rate or request a quota increase.

HTTP Status Code: 402

ThrottledException

The request was denied due to request throttling. Reduce the frequency of requests and try again.

HTTP Status Code: 429

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

ListMemoryRecords

Lists memory records in an AgentCore Memory resource based on specified criteria. We recommend using pagination to ensure that the operation returns quickly and successfully.

To use this operation, you must have the `bedrock-agentcore:ListMemoryRecords` permission.

Request Syntax

```
POST /memories/memoryId/memoryRecords HTTP/1.1
```

```
Content-type: application/json
```

```
{
  "maxResults": number,
  "memoryStrategyId": "string",
  "metadataFilters": [
    {
      "left": { ... },
      "operator": "string",
      "right": { ... }
    }
  ],
  "namespace": "string",
  "namespacePath": "string",
  "nextToken": "string"
}
```

URI Request Parameters

The request uses the following URI parameters.

memoryId

The identifier of the AgentCore Memory resource for which to list memory records.

Length Constraints: Minimum length of 12.

Pattern: `(arn:(aws|aws-cn|aws-us-gov):bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:memory/)?[a-zA-Z][a-zA-Z0-9-_{0,99}]-[a-zA-Z0-9]{10}`

Required: Yes

Request Body

The request accepts the following data in JSON format.

maxResults

The maximum number of results to return in a single call. The default value is 20.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

memoryStrategyId

The memory strategy identifier to filter memory records by. If specified, only memory records with this strategy ID are returned.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 100.

Pattern: [a-zA-Z0-9][a-zA-Z0-9-_*]

Required: No

metadataFilters

A list of metadata filter expressions to scope the returned memory records.

Type: Array of [MemoryMetadataFilterExpression](#) objects

Array Members: Minimum number of 1 item. Maximum number of 5 items.

Required: No

namespace

The namespace prefix to filter memory records by. Returns all memory records in namespaces that start with the provided prefix. Either namespace or namespacePath is required.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1024.

Pattern: `[a-zA-Z0-9/*][a-zA-Z0-9-_/]*(?::[a-zA-Z0-9-_/]*+)*[a-zA-Z0-9-_/]*`

Required: No

namespacePath

Use namespacePath for hierarchical retrievals. Return all memory records where namespace falls under the same parent hierarchy. Either namespace or namespacePath is required.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1024.

Pattern: `[a-zA-Z0-9/*][a-zA-Z0-9-_/]*(?::[a-zA-Z0-9-_/]*+)*[a-zA-Z0-9-_/]*`

Required: No

nextToken

The token for the next set of results. Use the value returned in the previous response in the next request to retrieve the next set of results.

Type: String

Required: No

Response Syntax

```
HTTP/1.1 200
Content-type: application/json

{
  "memoryRecordSummaries": [
    {
      "content": { ... },
      "createdAt": number,
      "memoryRecordId": "string",
      "memoryStrategyId": "string",
      "metadata": {
        "string" : { ... }
      }
    },
  ],
}
```

```
    "namespaces": [ "string" ],
    "score": number
  }
],
"nextToken": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

memoryRecordSummaries

The list of memory record summaries that match the specified criteria.

Type: Array of [MemoryRecordSummary](#) objects

nextToken

The token to use in a subsequent request to get the next set of results. This value is null when there are no more results to return.

Type: String

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

InvalidInputException

The input fails to satisfy the constraints specified by AgentCore. Check your input values and try again.

HTTP Status Code: 400

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ServiceException

The service encountered an internal error. Try your request again later.

HTTP Status Code: 500

ServiceQuotaExceededException

The exception that occurs when the request would cause a service quota to be exceeded. Review your service quotas and either reduce your request rate or request a quota increase.

HTTP Status Code: 402

ThrottledException

The request was denied due to request throttling. Reduce the frequency of requests and try again.

HTTP Status Code: 429

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)

- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

ListPaymentInstruments

List payment instruments for a manager.

Request Syntax

```
POST /payments/listPaymentInstruments HTTP/1.1
X-Amzn-Bedrock-AgentCore-Payments-User-Id: userId
X-Amzn-Bedrock-AgentCore-Payments-Agent-Name: agentName
Content-type: application/json

{
  "maxResults": number,
  "nextToken": "string",
  "paymentConnectorId": "string",
  "paymentManagerArn": "string"
}
```

URI Request Parameters

The request uses the following URI parameters.

agentName

The agent name associated with this request, used for observability.

Length Constraints: Minimum length of 0. Maximum length of 256.

userId

The user ID associated with the payment instruments.

Length Constraints: Minimum length of 0. Maximum length of 120.

Request Body

The request accepts the following data in JSON format.

maxResults

Maximum number of results to return in a single response.

Type: Integer

Required: No

nextToken

Token for pagination to retrieve the next set of results.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: `\S*`

Required: No

paymentConnectorId

The ID of the payment connector to filter by.

Type: String

Length Constraints: Minimum length of 12. Maximum length of 211.

Pattern: `([0-9a-z][-]?){1,100}-[0-9a-z]{10}`

Required: No

paymentManagerArn

The ARN of the payment manager that owns the payment instruments.

Type: String

Length Constraints: Minimum length of 66. Maximum length of 2048.

Pattern: `arn:(aws|aws-[a-z0-9-]+):bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:payment-manager/[a-z0-9]([a-z0-9-]{0,47}[a-z0-9])?-[a-z0-9]{10}`

Required: Yes

Response Syntax

```
HTTP/1.1 200
Content-type: application/json
```

```
{
  "nextToken": "string",
  "paymentInstruments": [
    {
      "createdAt": "string",
      "paymentConnectorId": "string",
      "paymentInstrumentId": "string",
      "paymentInstrumentType": "string",
      "paymentManagerArn": "string",
      "status": "string",
      "updatedAt": "string",
      "userId": "string"
    }
  ]
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

nextToken

Token for pagination to retrieve the next set of results.

Type: String

paymentInstruments

List of payment instrument summaries matching the request criteria.

Type: Array of [PaymentInstrumentSummary](#) objects

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

InternalServerErrorException

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)

- [AWS SDK for Ruby V3](#)

ListPaymentSessions

List payment sessions.

Request Syntax

```
POST /payments/listPaymentSessions HTTP/1.1
X-Amzn-Bedrock-AgentCore-Payments-User-Id: userId
X-Amzn-Bedrock-AgentCore-Payments-Agent-Name: agentName
Content-type: application/json
```

```
{
  "maxResults": number,
  "nextToken": "string",
  "paymentManagerArn": "string"
}
```

URI Request Parameters

The request uses the following URI parameters.

agentName

The agent name associated with this request, used for observability.

Length Constraints: Minimum length of 0. Maximum length of 256.

userId

The user ID associated with the payment sessions.

Length Constraints: Minimum length of 0. Maximum length of 120.

Request Body

The request accepts the following data in JSON format.

maxResults

Maximum number of results to return in a single response.

Type: Integer

Required: No

nextToken

Token for pagination to retrieve the next set of results.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: `\S*`

Required: No

paymentManagerArn

The ARN of the payment manager that owns the sessions.

Type: String

Length Constraints: Minimum length of 66. Maximum length of 2048.

Pattern: `arn:(aws|aws-[a-z0-9-]+):bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:payment-manager/[a-z0-9]([a-z0-9-]{0,47}[a-z0-9])?-[a-z0-9]{10}`

Required: Yes

Response Syntax

```
HTTP/1.1 200
Content-type: application/json

{
  "nextToken": "string",
  "paymentSessions": [
    {
      "createdAt": "string",
      "expiryTimeInMinutes": number,
      "paymentManagerArn": "string",
      "paymentSessionId": "string",
      "updatedAt": "string",
      "userId": "string"
    }
  ]
}
```

```
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

[nextToken](#)

Token for pagination to retrieve the next set of results.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: `\S*`

[paymentSessions](#)

List of payment session summaries matching the request criteria.

Type: Array of [PaymentSessionSummary](#) objects

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

InternalServerError

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

ListRecommendations

Lists all recommendations in the account, with optional filtering by status.

Request Syntax

```
GET /recommendations?maxResults=maxResults&nextToken=nextToken&status=statusFilter
HTTP/1.1
```

URI Request Parameters

The request uses the following URI parameters.

maxResults

The maximum number of results to return in the response. If the total number of results is greater than this value, use the token returned in the response in the `nextToken` field when making another request to return the next batch of results.

Valid Range: Minimum value of 1. Maximum value of 100.

nextToken

If the total number of results is greater than the `maxResults` value provided in the request, enter the token returned in the `nextToken` field in the response in this field to return the next batch of results.

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: `\S*`

statusFilter

Optional filter to return only recommendations with the specified status.

Valid Values: PENDING | IN_PROGRESS | COMPLETED | FAILED | DELETING

Request Body

The request does not have a request body.

Response Syntax

```
HTTP/1.1 200
Content-type: application/json

{
  "nextToken": "string",
  "recommendationSummaries": [
    {
      "createdAt": "string",
      "description": "string",
      "name": "string",
      "recommendationArn": "string",
      "recommendationId": "string",
      "status": "string",
      "type": "string",
      "updatedAt": "string"
    }
  ]
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

nextToken

If the total number of results is greater than the `maxResults` value provided in the request, use this token when making another request in the `nextToken` field to return the next batch of results.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: `\S*`

recommendationSummaries

The list of recommendation summaries.

Type: Array of [RecommendationSummary](#) objects

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

InternalServerError

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

ListSessions

Lists sessions in an AgentCore Memory resource based on specified criteria. We recommend using pagination to ensure that the operation returns quickly and successfully.

Empty sessions are automatically deleted after one day.

To use this operation, you must have the `bedrock-agentcore:ListSessions` permission.

Request Syntax

```
POST /memories/memoryId/actor/actorId/sessions HTTP/1.1
Content-type: application/json
```

```
{
  "filter": {
    "eventFilter": "string"
  },
  "maxResults": number,
  "nextToken": "string"
}
```

URI Request Parameters

The request uses the following URI parameters.

actorId

The identifier of the actor for which to list sessions.

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[a-zA-Z0-9][a-zA-Z0-9-_/]*(?::[a-zA-Z0-9-_/]+)*[a-zA-Z0-9-_/]*`

Required: Yes

memoryId

The identifier of the AgentCore Memory resource for which to list sessions.

Length Constraints: Minimum length of 12.

Pattern: (arn:(aws|aws-cn|aws-us-gov):bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:memory/)?[a-zA-Z][a-zA-Z0-9-_{0,99}-[a-zA-Z0-9]{10}

Required: Yes

Request Body

The request accepts the following data in JSON format.

filter

Filter criteria to apply when listing sessions.

Type: [SessionFilter](#) object

Required: No

maxResults

The maximum number of results to return in a single call. The default value is 20.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

nextToken

The token for the next set of results. Use the value returned in the previous response in the next request to retrieve the next set of results.

Type: String

Required: No

Response Syntax

```
HTTP/1.1 200
Content-type: application/json

{
  "nextToken": "string",
```

```
"sessionSummaries": [  
  {  
    "actorId": "string",  
    "createdAt": number,  
    "sessionId": "string"  
  }  
]
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

[nextToken](#)

The token to use in a subsequent request to get the next set of results. This value is null when there are no more results to return.

Type: String

[sessionSummaries](#)

The list of session summaries that match the specified criteria.

Type: Array of [SessionSummary](#) objects

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

InvalidInputException

The input fails to satisfy the constraints specified by AgentCore. Check your input values and try again.

HTTP Status Code: 400

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ServiceException

The service encountered an internal error. Try your request again later.

HTTP Status Code: 500

ServiceQuotaExceededException

The exception that occurs when the request would cause a service quota to be exceeded. Review your service quotas and either reduce your request rate or request a quota increase.

HTTP Status Code: 402

ThrottledException

The request was denied due to request throttling. Reduce the frequency of requests and try again.

HTTP Status Code: 429

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)

- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

ProcessPayment

Processes a payment using a payment instrument within a payment session.

Request Syntax

```
POST /payments/processPayment HTTP/1.1
X-Amzn-Bedrock-AgentCore-Payments-User-Id: userId
X-Amzn-Bedrock-AgentCore-Payments-Agent-Name: agentName
Content-type: application/json

{
  "clientToken": "string",
  "paymentInput": { ... },
  "paymentInstrumentId": "string",
  "paymentManagerArn": "string",
  "paymentSessionId": "string",
  "paymentType": "string"
}
```

URI Request Parameters

The request uses the following URI parameters.

agentName

The agent name associated with this request, used for observability.

Length Constraints: Minimum length of 0. Maximum length of 256.

userId

The user ID associated with this payment.

Length Constraints: Minimum length of 0. Maximum length of 120.

Request Body

The request accepts the following data in JSON format.

clientToken

A unique, case-sensitive identifier that you provide to ensure the idempotency of the request.

Type: String

Length Constraints: Minimum length of 33. Maximum length of 256.

Pattern: `[a-zA-Z0-9](-*[a-zA-Z0-9]){0,256}`

Required: No

paymentInput

The payment input details specific to the payment type.

Type: [PaymentInput](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: Yes

paymentInstrumentId

The ID of the payment instrument to use.

Type: String

Length Constraints: Fixed length of 34.

Pattern: `payment-instrument-[0-9a-zA-Z-]{15}`

Required: Yes

paymentManagerArn

The ARN of the payment manager.

Type: String

Length Constraints: Minimum length of 66. Maximum length of 2048.

Pattern: `arn:(aws|aws-[a-z0-9-]+):bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:payment-manager/[a-z0-9]([a-z0-9-]{0,47}[a-z0-9])?-[a-z0-9]{10}`

Required: Yes

paymentSessionId

The ID of the payment session.

Type: String

Length Constraints: Fixed length of 31.

Pattern: `payment-session-[0-9a-zA-Z-]{15}`

Required: Yes

paymentType

The type of payment to process.

Type: String

Valid Values: CRYPTO_X402

Required: Yes

Response Syntax

```
HTTP/1.1 200
Content-type: application/json

{
  "createdAt": "string",
  "paymentInstrumentId": "string",
  "paymentManagerArn": "string",
  "paymentOutput": { ... },
  "paymentSessionId": "string",
  "paymentType": "string",
  "processPaymentId": "string",
  "status": "string",
  "updatedAt": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

createdAt

The timestamp when the payment was created.

Type: Timestamp

paymentInstrumentId

The ID of the payment instrument used.

Type: String

Length Constraints: Fixed length of 34.

Pattern: `payment-instrument-[0-9a-zA-Z-]{15}`

paymentManagerArn

The ARN of the payment manager.

Type: String

Length Constraints: Minimum length of 66. Maximum length of 2048.

Pattern: `arn:(aws|aws-[a-z0-9-]+):bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:payment-manager/[a-z0-9]([a-z0-9-]{0,47}[a-z0-9])?-[a-z0-9]{10}`

paymentOutput

The payment output details specific to the payment type.

Type: [PaymentOutput](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

paymentSessionId

The ID of the payment session used.

Type: String

Length Constraints: Fixed length of 31.

Pattern: `payment-session-[0-9a-zA-Z-]{15}`

paymentType

The type of payment processed.

Type: String

Valid Values: CRYPTO_X402

processPaymentId

The unique identifier of the processed payment.

Type: String

Length Constraints: Fixed length of 36.

Pattern: [a-fA-F0-9]{8}-[a-fA-F0-9]{4}-[a-fA-F0-9]{4}-[a-fA-F0-9]{4}-[a-fA-F0-9]{12}

status

The status of the payment.

Type: String

Valid Values: PROOF_GENERATED

updatedAt

The timestamp when the payment was last updated.

Type: Timestamp

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

ConflictException

The exception that occurs when the request conflicts with the current state of the resource. This can happen when trying to modify a resource that is currently being modified by another request, or when trying to create a resource that already exists.

HTTP Status Code: 409

InternalServerErrorException

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ServiceQuotaExceededException

The exception that occurs when the request would cause a service quota to be exceeded. Review your service quotas and either reduce your request rate or request a quota increase.

HTTP Status Code: 402

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)

- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

RetrieveMemoryRecords

Searches for and retrieves memory records from an AgentCore Memory resource based on specified search criteria. We recommend using pagination to ensure that the operation returns quickly and successfully.

To use this operation, you must have the `bedrock-agentcore:RetrieveMemoryRecords` permission.

Request Syntax

```
POST /memories/memoryId/retrieve HTTP/1.1
Content-type: application/json
```

```
{
  "maxResults": number,
  "namespace": "string",
  "namespacePath": "string",
  "nextToken": "string",
  "searchCriteria": {
    "memoryStrategyId": "string",
    "metadataFilters": [
      {
        "left": { ... },
        "operator": "string",
        "right": { ... }
      }
    ],
    "searchQuery": "string",
    "topK": number
  }
}
```

URI Request Parameters

The request uses the following URI parameters.

memoryId

The identifier of the AgentCore Memory resource from which to retrieve memory records.

Length Constraints: Minimum length of 12.

Pattern: `(arn:(aws|aws-cn|aws-us-gov):bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:memory/)?[a-zA-Z][a-zA-Z0-9-_{0,99}-[a-zA-Z0-9]{10}`

Required: Yes

Request Body

The request accepts the following data in JSON format.

maxResults

The maximum number of results to return in a single call. The default value is 20.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

namespace

The namespace prefix to filter memory records by. Searches for memory records in namespaces that start with the provided prefix. Either namespace or namespacePath is required.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1024.

Pattern: `[a-zA-Z0-9/*][a-zA-Z0-9-_/]*([?:[a-zA-Z0-9-_/]*+)*[a-zA-Z0-9-_/]*]`

Required: No

namespacePath

Use namespacePath for hierarchical retrievals. Return all memory records where namespace falls under the same parent hierarchy. Either namespace or namespacePath is required.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1024.

Pattern: `[a-zA-Z0-9/*][a-zA-Z0-9-_/]**(?:[a-zA-Z0-9-_/]*+)*[a-zA-Z0-9-_/]**`

Required: No

[nextToken](#)

The token for the next set of results. Use the value returned in the previous response in the next request to retrieve the next set of results.

Type: String

Required: No

[searchCriteria](#)

The search criteria to use for finding relevant memory records. This includes the search query, memory strategy ID, and other search parameters.

Type: [SearchCriteria](#) object

Required: Yes

Response Syntax

```
HTTP/1.1 200
Content-type: application/json

{
  "memoryRecordSummaries": [
    {
      "content": { ... },
      "createdAt": number,
      "memoryRecordId": "string",
      "memoryStrategyId": "string",
      "metadata": {
        "string": { ... }
      },
      "namespaces": [ "string" ],
      "score": number
    }
  ],
  "nextToken": "string"
}
```

```
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

memoryRecordSummaries

The list of memory record summaries that match the search criteria, ordered by relevance.

Type: Array of [MemoryRecordSummary](#) objects

nextToken

The token to use in a subsequent request to get the next set of results. This value is null when there are no more results to return.

Type: String

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

InvalidInputException

The input fails to satisfy the constraints specified by AgentCore. Check your input values and try again.

HTTP Status Code: 400

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ServiceException

The service encountered an internal error. Try your request again later.

HTTP Status Code: 500

ServiceQuotaExceededException

The exception that occurs when the request would cause a service quota to be exceeded. Review your service quotas and either reduce your request rate or request a quota increase.

HTTP Status Code: 402

ThrottledException

The request was denied due to request throttling. Reduce the frequency of requests and try again.

HTTP Status Code: 429

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)

- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

SaveBrowserSessionProfile

Saves the current state of a browser session as a reusable profile in Amazon Bedrock AgentCore. A browser profile captures persistent browser data such as cookies and local storage from an active session, enabling you to reuse this data in future browser sessions.

To save a browser session profile, you must specify the profile identifier, browser identifier, and session ID. The session must be active when saving the profile. Once saved, the profile can be used with the `StartBrowserSession` operation to initialize new sessions with the stored browser state.

Browser profiles are useful for scenarios that require persistent authentication, maintaining user preferences across sessions, or continuing tasks that depend on previously stored browser data.

The following operations are related to `SaveBrowserSessionProfile`:

- [StartBrowserSession](#)
- [GetBrowserSession](#)

Request Syntax

```
PUT /browser-profiles/profileIdentifier/save HTTP/1.1
X-Amzn-Trace-Id: traceId
traceparent: traceParent
Content-type: application/json

{
  "browserIdentifier": "string",
  "clientToken": "string",
  "sessionId": "string"
}
```

URI Request Parameters

The request uses the following URI parameters.

profileIdentifier

The unique identifier for the browser profile. This identifier is used to reference the profile when starting new browser sessions. The identifier must follow the pattern of an alphanumeric name (up to 48 characters) followed by a hyphen and a 10-character alphanumeric suffix.

Pattern: `[a-zA-Z][a-zA-Z0-9_]{0,47}-[a-zA-Z0-9]{10}`

Required: Yes

traceId

The trace identifier for request tracking.

Length Constraints: Minimum length of 0. Maximum length of 1024.

traceParent

The parent trace information for distributed tracing.

Length Constraints: Minimum length of 0. Maximum length of 1024.

Request Body

The request accepts the following data in JSON format.

browserIdentifier

The unique identifier of the browser associated with the session from which to save the profile.

Type: String

Required: Yes

clientToken

A unique, case-sensitive identifier to ensure that the API request completes no more than one time. If this token matches a previous request, Amazon Bedrock AgentCore ignores the request, but does not return an error.

Type: String

Length Constraints: Minimum length of 33. Maximum length of 256.

Pattern: `[a-zA-Z0-9](-*[a-zA-Z0-9]){0,256}`

Required: No

sessionId

The unique identifier of the browser session from which to save the profile. The session must be active when saving the profile.

Type: String

Pattern: `[0-9a-zA-Z]{1,40}`

Required: Yes

Response Syntax

```
HTTP/1.1 200
Content-type: application/json

{
  "browserIdentifier": "string",
  "lastUpdatedAt": "string",
  "profileIdentifier": "string",
  "sessionId": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

browserIdentifier

The unique identifier of the browser associated with the session from which the profile was saved.

Type: String

lastUpdatedAt

The timestamp when the browser profile was last updated. This value is in ISO 8601 format.

Type: Timestamp

profileIdentifier

The unique identifier of the saved browser profile.

Type: String

Pattern: `[a-zA-Z][a-zA-Z0-9_]{0,47}-[a-zA-Z0-9]{10}`

sessionId

The unique identifier of the browser session from which the profile was saved.

Type: String

Pattern: `[0-9a-zA-Z]{1,40}`

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

ConflictException

The exception that occurs when the request conflicts with the current state of the resource. This can happen when trying to modify a resource that is currently being modified by another request, or when trying to create a resource that already exists.

HTTP Status Code: 409

InternalServerError

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)

- [AWS SDK for Ruby V3](#)

SearchRegistryRecords

Searches for registry records using semantic, lexical, or hybrid queries. Returns metadata for matching records ordered by relevance within the specified registry.

Request Syntax

```
POST /registry-records/search HTTP/1.1
Content-type: application/json
```

```
{
  "filters": JSON value,
  "maxResults": number,
  "registryIds": [ "string" ],
  "searchQuery": "string"
}
```

URI Request Parameters

The request does not use any URI parameters.

Request Body

The request accepts the following data in JSON format.

filters

A metadata filter expression to narrow search results. Uses structured JSON operators including field-level operators (`$eq`, `$ne`, `$in`) and logical operators (`$and`, `$or`) on filterable fields (`name`, `descriptorType`, `version`). For example, to filter by descriptor type: `{"descriptorType": {"$eq": "MCP"}}`. To combine filters: `{"$and": [{"descriptorType": {"$eq": "MCP"}}, {"name": {"$eq": "my-tool"}}]}`.

Type: JSON value

Required: No

maxResults

The maximum number of records to return in a single call. Valid values are 1 through 20. The default value is 10.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 20.

Required: No

registryIds

The list of registry identifiers to search within. Currently, you can specify exactly one registry identifier. You can provide either the full AWS Resource Name (ARN) or the 12-character alphanumeric registry ID.

Type: Array of strings

Array Members: Fixed number of 1 item.

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: `(arn:aws(-[^\:]+)?):bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:registry/)?[a-zA-Z0-9]{12,16}`

Required: Yes

searchQuery

The search query to find matching registry records.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Required: Yes

Response Syntax

```
HTTP/1.1 200
Content-type: application/json

{
  "registryRecords": [
    {
      "createdAt": "string",
      "description": "string",
      "descriptors": {
        "a2a": {
```

```
    "agentCard": {
      "inlineContent": "string",
      "schemaVersion": "string"
    },
    "agentSkills": {
      "skillDefinition": {
        "inlineContent": "string",
        "schemaVersion": "string"
      },
      "skillMd": {
        "inlineContent": "string"
      }
    },
    "custom": {
      "inlineContent": "string"
    },
    "mcp": {
      "server": {
        "inlineContent": "string",
        "schemaVersion": "string"
      },
      "tools": {
        "inlineContent": "string",
        "protocolVersion": "string"
      }
    }
  },
  "descriptorType": "string",
  "name": "string",
  "recordArn": "string",
  "recordId": "string",
  "registryArn": "string",
  "status": "string",
  "updatedAt": "string",
  "version": "string"
}
]
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

registryRecords

The list of registry records that match the search query, ordered by relevance.

Type: Array of [RegistryRecordSummary](#) objects

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

InternalServerError

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

UnauthorizedException

This exception is thrown when the JWT bearer token is invalid or not found for OAuth bearer token based access

HTTP Status Code: 401

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

StartBatchEvaluation

Starts a batch evaluation job that evaluates agent performance across multiple sessions. Batch evaluations pull agent traces from CloudWatch Logs or an existing online evaluation configuration and run specified evaluators and insights against them.

Request Syntax

```
POST /evaluations/batch-evaluate HTTP/1.1
```

```
Content-type: application/json
```

```
{
  "batchEvaluationName": "string",
  "clientToken": "string",
  "dataSourceConfig": { ... },
  "description": "string",
  "evaluationMetadata": { ... },
  "evaluators": [
    {
      "evaluatorId": "string"
    }
  ],
  "insights": [
    {
      "insightId": "string"
    }
  ],
  "kmsKeyArn": "string",
  "tags": {
    "string" : "string"
  }
}
```

URI Request Parameters

The request does not use any URI parameters.

Request Body

The request accepts the following data in JSON format.

batchEvaluationName

The name of the batch evaluation. Must be unique within your account.

Type: String

Pattern: `[a-zA-Z][a-zA-Z0-9_]{0,47}`

Required: Yes

clientToken

A unique, case-sensitive identifier to ensure that the API request completes no more than one time. If this token matches a previous request, the service ignores the request, but does not return an error.

Type: String

Length Constraints: Minimum length of 33. Maximum length of 256.

Pattern: `[a-zA-Z0-9](-*[a-zA-Z0-9]){0,256}`

Required: No

dataSourceConfig

The data source configuration that specifies where to pull agent session traces from for evaluation.

Type: [DataSourceConfig](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: Yes

description

The description of the batch evaluation.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 200.

Required: No

evaluationMetadata

Optional metadata for the evaluation, including session-specific ground truth data and test scenario identifiers.

Type: [EvaluationMetadata](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: No

evaluators

The list of evaluators to apply during the batch evaluation. Can include both built-in evaluators and custom evaluators. Maximum of 10 evaluators.

Type: Array of [Evaluator](#) objects

Array Members: Minimum number of 0 items. Maximum number of 10 items.

Required: No

insights

The list of insight analyses to run against sessions during the batch evaluation. Maximum of 10 insights.

Type: Array of [Insight](#) objects

Array Members: Minimum number of 0 items. Maximum number of 10 items.

Required: No

kmsKeyArn

The ARN of the AWS KMS key used to encrypt evaluation data. If provided, customer data is encrypted at rest with the specified key.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: `arn:aws(|-cn|-us-gov):kms:[a-zA-Z0-9-]*:[0-9]{12}:key/[a-zA-Z0-9-]{36}`

Required: No

tags

A map of tag keys and values to associate with the batch evaluation.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 50 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Key Pattern: `[a-zA-Z0-9\s._:/+=@-]*`

Value Length Constraints: Minimum length of 0. Maximum length of 256.

Value Pattern: `[a-zA-Z0-9\s._:/+=@-]*`

Required: No

Response Syntax

```
HTTP/1.1 202
```

```
Content-type: application/json
```

```
{
  "batchEvaluationArn": "string",
  "batchEvaluationId": "string",
  "batchEvaluationName": "string",
  "createdAt": "string",
  "description": "string",
  "evaluators": [
    {
      "evaluatorId": "string"
    }
  ],
  "insights": [
    {
      "insightId": "string"
    }
  ],
  "kmsKeyArn": "string",
  "outputConfig": { ... },
```

```
"status": "string",
"tags": {
  "string" : "string"
}
}
```

Response Elements

If the action is successful, the service sends back an HTTP 202 response.

The following data is returned in JSON format by the service.

batchEvaluationArn

The Amazon Resource Name (ARN) of the created batch evaluation.

Type: String

batchEvaluationId

The unique identifier of the created batch evaluation.

Type: String

Pattern: [a-zA-Z][a-zA-Z0-9-_{0,99}-[a-zA-Z0-9]{10}

batchEvaluationName

The name of the batch evaluation.

Type: String

Pattern: [a-zA-Z][a-zA-Z0-9_]{0,47}

createdAt

The timestamp when the batch evaluation was created.

Type: Timestamp

description

The description of the batch evaluation.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 200.

evaluators

The list of evaluators applied during the batch evaluation.

Type: Array of [Evaluator](#) objects

insights

The list of insight analyses applied during the batch evaluation.

Type: Array of [Insight](#) objects

Array Members: Minimum number of 0 items. Maximum number of 10 items.

kmsKeyArn

The ARN of the AWS KMS key used to encrypt evaluation data.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: `arn:aws(|-cn|-us-gov):kms:[a-zA-Z0-9-]*:[0-9]{12}:key/[a-zA-Z0-9-]{36}`

outputConfig

The output configuration specifying where evaluation results are written.

Type: [OutputConfig](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

status

The status of the batch evaluation.

Type: String

Valid Values: PENDING | IN_PROGRESS | COMPLETED | COMPLETED_WITH_ERRORS | FAILED | STOPPING | STOPPED | DELETING

tags

The tags associated with the batch evaluation.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 50 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Key Pattern: [a-zA-Z0-9\s._:/=+@-]*

Value Length Constraints: Minimum length of 0. Maximum length of 256.

Value Pattern: [a-zA-Z0-9\s._:/=+@-]*

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

ConflictException

The exception that occurs when the request conflicts with the current state of the resource. This can happen when trying to modify a resource that is currently being modified by another request, or when trying to create a resource that already exists.

HTTP Status Code: 409

InternalServerError

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ServiceQuotaExceededException

The exception that occurs when the request would cause a service quota to be exceeded. Review your service quotas and either reduce your request rate or request a quota increase.

HTTP Status Code: 402

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

UnauthorizedException

This exception is thrown when the JWT bearer token is invalid or not found for OAuth bearer token based access

HTTP Status Code: 401

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)

- [AWS SDK for Ruby V3](#)

StartBrowserSession

Creates and initializes a browser session in Amazon Bedrock AgentCore. The session enables agents to navigate and interact with web content, extract information from websites, and perform web-based tasks as part of their response generation.

To create a session, you must specify a browser identifier and a name. You can also configure the viewport dimensions to control the visible area of web content. The session remains active until it times out or you explicitly stop it using the `StopBrowserSession` operation.

The following operations are related to `StartBrowserSession`:

- [GetBrowserSession](#)
- [UpdateBrowserStream](#)
- [SaveBrowserSessionProfile](#)
- [StopBrowserSession](#)
- [InvokeBrowser](#)

Request Syntax

```
PUT /browsers/browserIdentifier/sessions/start HTTP/1.1
X-Amzn-Trace-Id: traceId
traceparent: traceParent
Content-type: application/json

{
  "certificates": [
    {
      "location": { ... }
    }
  ],
  "clientToken": "string",
  "enterprisePolicies": [
    {
      "location": { ... },
      "type": "string"
    }
  ],
  "extensions": [
```

```
{
  "location": { ... }
},
"name": "string",
"profileConfiguration": {
  "profileIdentifier": "string"
},
"proxyConfiguration": {
  "bypass": {
    "domainPatterns": [ "string" ]
  },
  "proxies": [
    { ... }
  ]
},
"sessionTimeoutSeconds": number,
"viewPort": {
  "height": number,
  "width": number
}
}
```

URI Request Parameters

The request uses the following URI parameters.

browserIdentifier

The unique identifier of the browser to use for this session. This identifier specifies which browser environment to initialize for the session.

Required: Yes

traceId

The trace identifier for request tracking.

Length Constraints: Minimum length of 0. Maximum length of 1024.

traceParent

The parent trace information for distributed tracing.

Length Constraints: Minimum length of 0. Maximum length of 1024.

Request Body

The request accepts the following data in JSON format.

certificates

A list of certificates to install in the browser session.

Type: Array of [Certificate](#) objects

Array Members: Minimum number of 1 item. Maximum number of 200 items.

Required: No

clientToken

A unique, case-sensitive identifier to ensure that the API request completes no more than one time. If this token matches a previous request, Amazon Bedrock AgentCore ignores the request, but does not return an error. This parameter helps prevent the creation of duplicate sessions if there are temporary network issues.

Type: String

Length Constraints: Minimum length of 33. Maximum length of 256.

Pattern: `[a-zA-Z0-9](-*[a-zA-Z0-9]){0,256}`

Required: No

enterprisePolicies

A list of files containing enterprise policies for the browser.

Type: Array of [BrowserEnterprisePolicy](#) objects

Array Members: Minimum number of 0 items. Maximum number of 100 items.

Required: No

extensions

A list of browser extensions to load into the browser session.

Type: Array of [BrowserExtension](#) objects

Array Members: Minimum number of 1 item. Maximum number of 10 items.

Required: No

name

The name of the browser session. This name helps you identify and manage the session. The name does not need to be unique.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 100.

Required: No

profileConfiguration

The browser profile configuration to use for this session. A browser profile contains persistent data such as cookies and local storage that can be reused across multiple browser sessions. If specified, the session initializes with the profile's stored data, enabling continuity for tasks that require authentication or personalized settings.

Type: [BrowserProfileConfiguration](#) object

Required: No

proxyConfiguration

Optional proxy configuration for routing browser traffic through customer-specified proxy servers. When provided, enables HTTP Basic authentication via AWS Secrets Manager and domain-based routing rules. Requires `secretsmanager:GetSecretValue` IAM permission for the specified secret ARNs.

Type: [ProxyConfiguration](#) object

Required: No

sessionTimeoutSeconds

The duration in seconds (time-to-live) after which the session automatically terminates, regardless of ongoing activity. Defaults to 3600 seconds (1 hour). Recommended minimum: 60 seconds. Maximum allowed: 28,800 seconds (8 hours).

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 28800.

Required: No

[viewPort](#)

The dimensions of the browser viewport for this session. This determines the visible area of the web content and affects how web pages are rendered. If not specified, Amazon Bedrock AgentCore uses a default viewport size.

Type: [ViewPort](#) object

Required: No

Response Syntax

```
HTTP/1.1 200
Content-type: application/json

{
  "browserIdentifier": "string",
  "createdAt": "string",
  "sessionId": "string",
  "streams": {
    "automationStream": {
      "streamEndpoint": "string",
      "streamStatus": "string"
    },
    "liveViewStream": {
      "streamEndpoint": "string"
    }
  }
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

[browserIdentifier](#)

The identifier of the browser.

Type: String

createdAt

The timestamp when the browser session was created.

Type: Timestamp

sessionId

The unique identifier of the created browser session.

Type: String

Pattern: `[0-9a-zA-Z]{1,40}`

streams

The streams associated with this browser session. These include the automation stream and live view stream.

Type: [BrowserSessionStream](#) object

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

ConflictException

The exception that occurs when the request conflicts with the current state of the resource. This can happen when trying to modify a resource that is currently being modified by another request, or when trying to create a resource that already exists.

HTTP Status Code: 409

InternalServerErrorException

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ServiceQuotaExceededException

The exception that occurs when the request would cause a service quota to be exceeded. Review your service quotas and either reduce your request rate or request a quota increase.

HTTP Status Code: 402

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)

- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

StartCodeInterpreterSession

Creates and initializes a code interpreter session in Amazon Bedrock AgentCore. The session enables agents to execute code as part of their response generation, supporting programming languages such as Python for data analysis, visualization, and computation tasks.

To create a session, you must specify a code interpreter identifier and a name. The session remains active until it times out or you explicitly stop it using the `StopCodeInterpreterSession` operation.

The following operations are related to `StartCodeInterpreterSession`:

- [InvokeCodeInterpreter](#)
- [GetCodeInterpreterSession](#)
- [StopCodeInterpreterSession](#)

Request Syntax

```
PUT /code-interpreters/codeInterpreterIdentifier/sessions/start HTTP/1.1
X-Amzn-Trace-Id: traceId
traceparent: traceParent
Content-type: application/json
```

```
{
  "certificates": [
    {
      "location": { ... }
    }
  ],
  "clientToken": "string",
  "name": "string",
  "sessionTimeoutSeconds": number
}
```

URI Request Parameters

The request uses the following URI parameters.

codeInterpreterIdentifier

The unique identifier of the code interpreter to use for this session. This identifier specifies which code interpreter environment to initialize for the session.

Required: Yes

traceId

The trace identifier for request tracking.

Length Constraints: Minimum length of 0. Maximum length of 1024.

traceParent

The parent trace information for distributed tracing.

Length Constraints: Minimum length of 0. Maximum length of 1024.

Request Body

The request accepts the following data in JSON format.

certificates

A list of certificates to install in the code interpreter session.

Type: Array of [Certificate](#) objects

Array Members: Minimum number of 1 item. Maximum number of 200 items.

Required: No

clientToken

A unique, case-sensitive identifier to ensure that the API request completes no more than one time. If this token matches a previous request, Amazon Bedrock AgentCore ignores the request, but does not return an error. This parameter helps prevent the creation of duplicate sessions if there are temporary network issues.

Type: String

Length Constraints: Minimum length of 33. Maximum length of 256.

Pattern: `[a-zA-Z0-9](-*[a-zA-Z0-9]){0,256}`

Required: No

name

The name of the code interpreter session. This name helps you identify and manage the session. The name does not need to be unique.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 100.

Required: No

sessionTimeoutSeconds

The duration in seconds (time-to-live) after which the session automatically terminates, regardless of ongoing activity. Defaults to 900 seconds (15 minutes). Recommended minimum: 60 seconds. Maximum allowed: 28,800 seconds (8 hours).

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 28800.

Required: No

Response Syntax

```
HTTP/1.1 200
Content-type: application/json

{
  "codeInterpreterIdentifier": "string",
  "createdAt": "string",
  "sessionId": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

codeInterpreterIdentifier

The identifier of the code interpreter.

Type: String

createdAt

The time at which the code interpreter session was created.

Type: Timestamp

sessionId

The unique identifier of the created code interpreter session.

Type: String

Pattern: `[0-9a-zA-Z]{1,40}`

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

ConflictException

The exception that occurs when the request conflicts with the current state of the resource. This can happen when trying to modify a resource that is currently being modified by another request, or when trying to create a resource that already exists.

HTTP Status Code: 409

InternalServerError

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ServiceQuotaExceededException

The exception that occurs when the request would cause a service quota to be exceeded. Review your service quotas and either reduce your request rate or request a quota increase.

HTTP Status Code: 402

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)

- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

StartMemoryExtractionJob

Starts a memory extraction job that processes events that failed extraction previously in an AgentCore Memory resource and produces structured memory records. When earlier extraction attempts have left events unprocessed, this job will pick up and extract those as well.

To use this operation, you must have the `bedrock-agentcore:StartMemoryExtractionJob` permission.

Request Syntax

```
POST /memories/memoryId/extractionJobs/start HTTP/1.1
Content-type: application/json
```

```
{
  "clientToken": "string",
  "extractionJob": {
    "jobId": "string"
  }
}
```

URI Request Parameters

The request uses the following URI parameters.

memoryId

The unique identifier of the memory for which to start extraction jobs.

Length Constraints: Minimum length of 12.

Pattern: `(arn:(aws|aws-cn|aws-us-gov):bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:memory/)?[a-zA-Z][a-zA-Z0-9-_{0,99}]-[a-zA-Z0-9]{10}`

Required: Yes

Request Body

The request accepts the following data in JSON format.

clientToken

A unique, case-sensitive identifier to ensure idempotent processing of the request.

Type: String

Required: No

extractionJob

Extraction job to start in this operation.

Type: [ExtractionJob](#) object

Required: Yes

Response Syntax

```
HTTP/1.1 200
Content-type: application/json

{
  "jobId": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

jobId

Extraction Job ID that was attempted to start.

Type: String

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ServiceException

The service encountered an internal error. Try your request again later.

HTTP Status Code: 500

ServiceQuotaExceededException

The exception that occurs when the request would cause a service quota to be exceeded. Review your service quotas and either reduce your request rate or request a quota increase.

HTTP Status Code: 402

ThrottledException

The request was denied due to request throttling. Reduce the frequency of requests and try again.

HTTP Status Code: 429

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

StartRecommendation

Starts a recommendation job that analyzes agent traces and generates optimization suggestions for system prompts or tool descriptions to improve agent performance.

Request Syntax

```
POST /recommendations HTTP/1.1
Content-type: application/json

{
  "clientToken": "string",
  "description": "string",
  "kmsKeyArn": "string",
  "name": "string",
  "recommendationConfig": { ... },
  "tags": {
    "string" : "string"
  },
  "type": "string"
}
```

URI Request Parameters

The request does not use any URI parameters.

Request Body

The request accepts the following data in JSON format.

clientToken

A unique, case-sensitive identifier to ensure that the API request completes no more than one time. If this token matches a previous request, the service ignores the request, but does not return an error.

Type: String

Length Constraints: Minimum length of 33. Maximum length of 256.

Pattern: `[a-zA-Z0-9](-*[a-zA-Z0-9]){0,256}`

Required: No

description

The description of the recommendation.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 4096.

Required: No

kmsKeyArn

The ARN of the AWS KMS key used to encrypt recommendation data. If provided, customer data is encrypted at rest with the specified key.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: `arn:aws(|-cn|-us-gov):kms:[a-zA-Z0-9-]*:[0-9]{12}:key/[a-zA-Z0-9-]{36}`

Required: No

name

The name of the recommendation. Must be unique within your account.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 100.

Pattern: `[a-zA-Z][a-zA-Z0-9_-]{0,47}`

Required: Yes

recommendationConfig

The configuration for the recommendation, including the input to optimize, agent traces to analyze, and evaluation settings.

Type: [RecommendationConfig](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: Yes

tags

A map of tag keys and values to associate with the recommendation.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 50 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Key Pattern: `[a-zA-Z0-9\s._:/=+@-]*`

Value Length Constraints: Minimum length of 0. Maximum length of 256.

Value Pattern: `[a-zA-Z0-9\s._:/=+@-]*`

Required: No

type

The type of recommendation to generate. Valid values are `SYSTEM_PROMPT_RECOMMENDATION` for system prompt optimization or `TOOL_DESCRIPTION_RECOMMENDATION` for tool description optimization.

Type: String

Valid Values: `SYSTEM_PROMPT_RECOMMENDATION` | `TOOL_DESCRIPTION_RECOMMENDATION`

Required: Yes

Response Syntax

```
HTTP/1.1 202
Content-type: application/json

{
  "createdAt": "string",
  "description": "string",
  "name": "string",
  "recommendationArn": "string",
  "recommendationConfig": { ... },
  "recommendationId": "string",
```

```
"status": "string",  
"type": "string",  
"updatedAt": "string"  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 202 response.

The following data is returned in JSON format by the service.

[createdAt](#)

The timestamp when the recommendation was created.

Type: Timestamp

[description](#)

The description of the recommendation.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 4096.

[name](#)

The name of the recommendation.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 100.

Pattern: `[a-zA-Z][a-zA-Z0-9_-]{0,47}`

[recommendationArn](#)

The Amazon Resource Name (ARN) of the created recommendation.

Type: String

Pattern: `arn:aws[a-zA-Z-]*:bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:recommendation/[0-9a-zA-Z_-]{1,48}-[0-9A-Z]{10}`

[recommendationConfig](#)

The configuration for the recommendation.

Type: [RecommendationConfig](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

[recommendationId](#)

The unique identifier of the created recommendation.

Type: String

Pattern: `[0-9a-zA-Z_-]{1,48}-[0-9A-Z]{10}`

[status](#)

The status of the recommendation.

Type: String

Valid Values: PENDING | IN_PROGRESS | COMPLETED | FAILED | DELETING

[type](#)

The type of recommendation.

Type: String

Valid Values: SYSTEM_PROMPT_RECOMMENDATION | TOOL_DESCRIPTION_RECOMMENDATION

[updatedAt](#)

The timestamp when the recommendation was last updated.

Type: Timestamp

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

ConflictException

The exception that occurs when the request conflicts with the current state of the resource. This can happen when trying to modify a resource that is currently being modified by another request, or when trying to create a resource that already exists.

HTTP Status Code: 409

InternalServerError

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ServiceQuotaExceededException

The exception that occurs when the request would cause a service quota to be exceeded. Review your service quotas and either reduce your request rate or request a quota increase.

HTTP Status Code: 402

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

StopBatchEvaluation

Stops a running batch evaluation. Sessions that have already been evaluated retain their results.

Request Syntax

```
POST /evaluations/batch-evaluate/batchEvaluationId/stop HTTP/1.1
```

URI Request Parameters

The request uses the following URI parameters.

[batchEvaluationId](#)

The unique identifier of the batch evaluation to stop.

Pattern: `[a-zA-Z][a-zA-Z0-9-_{0,99}-[a-zA-Z0-9]{10}`

Required: Yes

Request Body

The request does not have a request body.

Response Syntax

```
HTTP/1.1 202
Content-type: application/json

{
  "batchEvaluationArn": "string",
  "batchEvaluationId": "string",
  "description": "string",
  "status": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 202 response.

The following data is returned in JSON format by the service.

batchEvaluationArn

The Amazon Resource Name (ARN) of the stopped batch evaluation.

Type: String

batchEvaluationId

The unique identifier of the stopped batch evaluation.

Type: String

Pattern: `[a-zA-Z][a-zA-Z0-9-_{0,99}-[a-zA-Z0-9]{10}`

description

The description of the batch evaluation.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 200.

status

The status of the batch evaluation after the stop request.

Type: String

Valid Values: PENDING | IN_PROGRESS | COMPLETED | COMPLETED_WITH_ERRORS | FAILED | STOPPING | STOPPED | DELETING

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

ConflictException

The exception that occurs when the request conflicts with the current state of the resource. This can happen when trying to modify a resource that is currently being modified by another request, or when trying to create a resource that already exists.

HTTP Status Code: 409

InternalServerError

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

UnauthorizedException

This exception is thrown when the JWT bearer token is invalid or not found for OAuth bearer token based access

HTTP Status Code: 401

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

StopBrowserSession

Terminates an active browser session in Amazon Bedrock AgentCore. This operation stops the session, releases associated resources, and makes the session unavailable for further use.

To stop a browser session, you must specify both the browser identifier and the session ID. Once stopped, a session cannot be restarted; you must create a new session using `StartBrowserSession`.

The following operations are related to `StopBrowserSession`:

- [StartBrowserSession](#)
- [GetBrowserSession](#)

Request Syntax

```
PUT /browsers/browserIdentifier/sessions/stop?sessionId=sessionId HTTP/1.1
X-Amzn-Trace-Id: traceId
traceparent: traceParent
Content-type: application/json

{
  "clientToken": "string"
}
```

URI Request Parameters

The request uses the following URI parameters.

[browserIdentifier](#)

The unique identifier of the browser associated with the session.

Required: Yes

[sessionId](#)

The unique identifier of the browser session to stop.

Pattern: `[0-9a-zA-Z]{1,40}`

Required: Yes

traceId

The trace identifier for request tracking.

Length Constraints: Minimum length of 0. Maximum length of 1024.

traceParent

The parent trace information for distributed tracing.

Length Constraints: Minimum length of 0. Maximum length of 1024.

Request Body

The request accepts the following data in JSON format.

clientToken

A unique, case-sensitive identifier to ensure that the API request completes no more than one time. If this token matches a previous request, Amazon Bedrock AgentCore ignores the request, but does not return an error.

Type: String

Length Constraints: Minimum length of 33. Maximum length of 256.

Pattern: `[a-zA-Z0-9](-*[a-zA-Z0-9]){0,256}`

Required: No

Response Syntax

```
HTTP/1.1 200
Content-type: application/json

{
  "browserIdentifier": "string",
  "lastUpdatedAt": "string",
  "sessionId": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

browserIdentifier

The identifier of the browser.

Type: String

lastUpdatedAt

The time at which the browser session was last updated.

Type: Timestamp

sessionId

The identifier of the browser session.

Type: String

Pattern: `[0-9a-zA-Z]{1,40}`

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

ConflictException

The exception that occurs when the request conflicts with the current state of the resource. This can happen when trying to modify a resource that is currently being modified by another request, or when trying to create a resource that already exists.

HTTP Status Code: 409

InternalServerErrorException

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ServiceQuotaExceededException

The exception that occurs when the request would cause a service quota to be exceeded. Review your service quotas and either reduce your request rate or request a quota increase.

HTTP Status Code: 402

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

StopCodeInterpreterSession

Terminates an active code interpreter session in Amazon Bedrock AgentCore. This operation stops the session, releases associated resources, and makes the session unavailable for further use.

To stop a code interpreter session, you must specify both the code interpreter identifier and the session ID. Once stopped, a session cannot be restarted; you must create a new session using `StartCodeInterpreterSession`.

The following operations are related to `StopCodeInterpreterSession`:

- [StartCodeInterpreterSession](#)
- [GetCodeInterpreterSession](#)

Request Syntax

```
PUT /code-interpreters/codeInterpreterIdentifier/sessions/stop?sessionId=sessionId
HTTP/1.1
X-Amzn-Trace-Id: traceId
traceparent: traceParent
Content-type: application/json

{
  "clientToken": "string"
}
```

URI Request Parameters

The request uses the following URI parameters.

[codeInterpreterIdentifier](#)

The unique identifier of the code interpreter associated with the session.

Required: Yes

[sessionId](#)

The unique identifier of the code interpreter session to stop.

Pattern: `[0-9a-zA-Z]{1,40}`

Required: Yes

traceId

The trace identifier for request tracking.

Length Constraints: Minimum length of 0. Maximum length of 1024.

traceParent

The parent trace information for distributed tracing.

Length Constraints: Minimum length of 0. Maximum length of 1024.

Request Body

The request accepts the following data in JSON format.

clientToken

A unique, case-sensitive identifier to ensure that the API request completes no more than one time. If this token matches a previous request, Amazon Bedrock AgentCore ignores the request, but does not return an error.

Type: String

Length Constraints: Minimum length of 33. Maximum length of 256.

Pattern: `[a-zA-Z0-9](-*[a-zA-Z0-9]){0,256}`

Required: No

Response Syntax

```
HTTP/1.1 200
Content-type: application/json

{
  "codeInterpreterIdentifier": "string",
  "lastUpdatedAt": "string",
  "sessionId": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

codeInterpreterIdentifier

The identifier of the code interpreter.

Type: String

lastUpdatedAt

The timestamp when the code interpreter session was last updated.

Type: Timestamp

sessionId

The identifier of the code interpreter session.

Type: String

Pattern: `[0-9a-zA-Z]{1,40}`

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

ConflictException

The exception that occurs when the request conflicts with the current state of the resource. This can happen when trying to modify a resource that is currently being modified by another request, or when trying to create a resource that already exists.

HTTP Status Code: 409

InternalServerErrorException

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ServiceQuotaExceededException

The exception that occurs when the request would cause a service quota to be exceeded. Review your service quotas and either reduce your request rate or request a quota increase.

HTTP Status Code: 402

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

StopRuntimeSession

Stops a session that is running in an running AgentCore Runtime agent.

Request Syntax

```
POST /runtimes/agentRuntimeArn/stopruntimesession?qualifier=qualifier HTTP/1.1
X-Amzn-Bedrock-AgentCore-Runtime-Session-Id: runtimeSessionId
Content-type: application/json

{
  "clientToken": "string"
}
```

URI Request Parameters

The request uses the following URI parameters.

agentRuntimeArn

The ARN of the agent that contains the session that you want to stop.

Required: Yes

qualifier

Optional qualifier to specify an agent alias, such as `prodcode>` or `dev`. If you don't provide a value, the DEFAULT alias is used.

runtimeSessionId

The ID of the session that you want to stop.

Length Constraints: Minimum length of 33. Maximum length of 256.

Required: Yes

Request Body

The request accepts the following data in JSON format.

clientToken

Idempotent token used to identify the request. If you use the same token with multiple requests, the same response is returned. Use ClientToken to prevent the same request from being processed more than once.

Type: String

Length Constraints: Minimum length of 33. Maximum length of 256.

Pattern: `[a-zA-Z0-9](-*[a-zA-Z0-9]){0,256}`

Required: No

Response Syntax

```
HTTP/1.1 statusCode  
X-Amzn-Bedrock-AgentCore-Runtime-Session-Id: runtimeSessionId
```

Response Elements

If the action is successful, the service sends back the following HTTP response.

statusCode

The status code of the request to stop the session.

The response returns the following HTTP headers.

runtimeSessionId

The ID of the session that you requested to stop.

Length Constraints: Minimum length of 1. Maximum length of 100.

Pattern: `[a-zA-Z0-9][a-zA-Z0-9-_*]{0,99}`

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

ConflictException

The exception that occurs when the request conflicts with the current state of the resource. This can happen when trying to modify a resource that is currently being modified by another request, or when trying to create a resource that already exists.

HTTP Status Code: 409

InternalServerError

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

RetryableConflictException

The exception that occurs when there is a retryable conflict performing an operation. This is a temporary condition that may resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 409

RuntimeClientError

The exception that occurs when there is an error in the runtime client. This can happen due to network issues, invalid configuration, or other client-side problems. Check the error message for specific details about the error.

HTTP Status Code: 424

ServiceQuotaExceededException

The exception that occurs when the request would cause a service quota to be exceeded. Review your service quotas and either reduce your request rate or request a quota increase.

HTTP Status Code: 402

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

UnauthorizedException

This exception is thrown when the JWT bearer token is invalid or not found for OAuth bearer token based access

HTTP Status Code: 401

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)

- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

UpdateABTest

Updates an A/B test's configuration, including variants, traffic allocation, evaluation settings, or execution status.

Request Syntax

```
PUT /ab-tests/abTestId HTTP/1.1
Content-type: application/json

{
  "clientToken": "string",
  "description": "string",
  "evaluationConfig": { ... },
  "executionStatus": "string",
  "gatewayFilter": {
    "targetPaths": [ "string" ]
  },
  "name": "string",
  "roleArn": "string",
  "variants": [
    {
      "name": "string",
      "variantConfiguration": {
        "configurationBundle": {
          "bundleArn": "string",
          "bundleVersion": "string"
        },
        "target": {
          "name": "string"
        }
      },
      "weight": number
    }
  ]
}
```

URI Request Parameters

The request uses the following URI parameters.

abTestId

The unique identifier of the A/B test to update.

Pattern: `[a-zA-Z][a-zA-Z0-9-_{0,99}-[a-zA-Z0-9]{10}`

Required: Yes

Request Body

The request accepts the following data in JSON format.

clientToken

A unique, case-sensitive identifier to ensure that the API request completes no more than one time. If this token matches a previous request, the service ignores the request, but does not return an error.

Type: String

Length Constraints: Minimum length of 33. Maximum length of 256.

Pattern: `[a-zA-Z0-9](-*[a-zA-Z0-9]){0,256}`

Required: No

description

The updated description of the A/B test.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 200.

Required: No

evaluationConfig

The updated evaluation configuration.

Type: [ABTestEvaluationConfig](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: No

executionStatus

The updated execution status to enable or disable the A/B test.

Type: String

Valid Values: PAUSED | RUNNING | STOPPED | NOT_STARTED

Required: No

gatewayFilter

The updated gateway filter.

Type: [GatewayFilter](#) object

Required: No

name

The updated name of the A/B test.

Type: String

Pattern: [a-zA-Z][a-zA-Z0-9_]{0,47}

Required: No

roleArn

The updated IAM role ARN.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: arn:aws(-[^:]+)?:iam::([0-9]{12})?:role/.+

Required: No

variants

The updated list of variants.

Type: Array of [Variant](#) objects

Array Members: Fixed number of 2 items.

Required: No

Response Syntax

```
HTTP/1.1 202
Content-type: application/json

{
  "abTestArn": "string",
  "abTestId": "string",
  "executionStatus": "string",
  "status": "string",
  "updatedAt": number
}
```

Response Elements

If the action is successful, the service sends back an HTTP 202 response.

The following data is returned in JSON format by the service.

[abTestArn](#)

The Amazon Resource Name (ARN) of the updated A/B test.

Type: String

Pattern: `arn:aws[a-zA-Z-]*:bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:ab-test/[a-zA-Z][a-zA-Z0-9-]{0,99}-[a-zA-Z0-9]{10}`

[abTestId](#)

The unique identifier of the updated A/B test.

Type: String

Pattern: `[a-zA-Z][a-zA-Z0-9-]{0,99}-[a-zA-Z0-9]{10}`

[executionStatus](#)

The execution status of the A/B test.

Type: String

Valid Values: PAUSED | RUNNING | STOPPED | NOT_STARTED

status

The status of the A/B test.

Type: String

Valid Values: CREATING | ACTIVE | CREATE_FAILED | UPDATING | UPDATE_FAILED | DELETING | DELETE_FAILED | FAILED

updatedAt

The timestamp when the A/B test was updated.

Type: Timestamp

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

ConflictException

The exception that occurs when the request conflicts with the current state of the resource. This can happen when trying to modify a resource that is currently being modified by another request, or when trying to create a resource that already exists.

HTTP Status Code: 409

InternalServerError

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ServiceQuotaExceededException

The exception that occurs when the request would cause a service quota to be exceeded. Review your service quotas and either reduce your request rate or request a quota increase.

HTTP Status Code: 402

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

UnauthorizedException

This exception is thrown when the JWT bearer token is invalid or not found for OAuth bearer token based access

HTTP Status Code: 401

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

UpdateBrowserStream

Updates a browser stream. To use this operation, you must have permissions to perform the `bedrock:UpdateBrowserStream` action.

Request Syntax

```
PUT /browsers/browserIdentifier/sessions/streams/update?sessionId=sessionId HTTP/1.1  
Content-type: application/json
```

```
{  
  "clientToken": "string",  
  "streamUpdate": { ... }  
}
```

URI Request Parameters

The request uses the following URI parameters.

browserIdentifier

The identifier of the browser.

Required: Yes

sessionId

The identifier of the browser session.

Pattern: `[0-9a-zA-Z]{1,40}`

Required: Yes

Request Body

The request accepts the following data in JSON format.

clientToken

A unique, case-sensitive identifier to ensure that the operation completes no more than one time. If this token matches a previous request, Amazon Bedrock ignores the request, but does not return an error.

Type: String

Length Constraints: Minimum length of 33. Maximum length of 256.

Pattern: `[a-zA-Z0-9](-*[a-zA-Z0-9]){0,256}`

Required: No

streamUpdate

The update to apply to the browser stream.

Type: [StreamUpdate](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: Yes

Response Syntax

```
HTTP/1.1 200
Content-type: application/json

{
  "browserIdentifier": "string",
  "sessionId": "string",
  "streams": {
    "automationStream": {
      "streamEndpoint": "string",
      "streamStatus": "string"
    },
    "liveViewStream": {
      "streamEndpoint": "string"
    }
  },
  "updatedAt": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

browserIdentifier

The identifier of the browser.

Type: String

sessionId

The identifier of the browser session.

Type: String

Pattern: `[0-9a-zA-Z]{1,40}`

streams

The collection of streams associated with a browser session in Amazon Bedrock AgentCore. These streams provide different ways to interact with and observe the browser session, including programmatic control and visual representation of the browser content.

Type: [BrowserSessionStream](#) object

updatedAt

The time at which the browser stream was updated.

Type: Timestamp

Errors

For information about the errors that are common to all actions, see [Common Error Types](#).

AccessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

HTTP Status Code: 403

ConflictException

The exception that occurs when the request conflicts with the current state of the resource. This can happen when trying to modify a resource that is currently being modified by another request, or when trying to create a resource that already exists.

HTTP Status Code: 409

InternalServerErrorException

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

HTTP Status Code: 500

ResourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

HTTP Status Code: 404

ServiceQuotaExceededException

The exception that occurs when the request would cause a service quota to be exceeded. Review your service quotas and either reduce your request rate or request a quota increase.

HTTP Status Code: 402

ThrottlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

HTTP Status Code: 429

ValidationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET V4](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

Data Types

The Amazon Bedrock AgentCore API contains several data types that various actions use. This section describes each data type in detail.

Note

The order of each element in a data type structure is not guaranteed. Applications should not assume a particular order.

The following data types are supported:

- [A2aDescriptor](#)
- [ABTestEvaluationConfig](#)
- [ABTestResults](#)
- [ABTestSummary](#)
- [ActorSummary](#)
- [AffectedSession](#)
- [AgentCardDefinition](#)
- [AgentSkillsDescriptor](#)
- [AgentTracesConfig](#)
- [Amount](#)
- [AutomationStream](#)
- [AutomationStreamUpdate](#)
- [AvailableLimits](#)
- [BasicAuth](#)
- [BatchEvaluationSummary](#)
- [BatchEvaluationTraceConfig](#)
- [Branch](#)
- [BranchFilter](#)
- [BrowserAction](#)
- [BrowserActionResult](#)

- [BrowserEnterprisePolicy](#)
- [BrowserExtension](#)
- [BrowserProfileConfiguration](#)
- [BrowserSessionStream](#)
- [BrowserSessionSummary](#)
- [Certificate](#)
- [CertificateLocation](#)
- [CloudWatchFilterConfig](#)
- [CloudWatchLogsFilter](#)
- [CloudWatchLogsRule](#)
- [CloudWatchLogsSource](#)
- [CloudWatchLogsTraceConfig](#)
- [CloudWatchOutputConfig](#)
- [CodeInterpreterResult](#)
- [CodeInterpreterSessionSummary](#)
- [CodeInterpreterStreamOutput](#)
- [CoinbaseCdpTokenRequestInput](#)
- [CoinbaseCdpTokenResponseOutput](#)
- [ConfidenceInterval](#)
- [ConfigurationBundleRef](#)
- [ConfigurationBundleToolEntry](#)
- [Content](#)
- [ContentBlock](#)
- [ContentDeltaEvent](#)
- [ContentStartEvent](#)
- [ContentStopEvent](#)
- [Context](#)
- [ControlStats](#)
- [Conversational](#)
- [CryptoX402PaymentInput](#)

- [CryptoX402PaymentOutput](#)
- [CustomDescriptor](#)
- [DataSourceConfig](#)
- [Descriptors](#)
- [EmbeddedCryptoWallet](#)
- [EvaluationContent](#)
- [EvaluationExpectedTrajectory](#)
- [EvaluationInput](#)
- [EvaluationJobResults](#)
- [EvaluationMetadata](#)
- [EvaluationReferenceInput](#)
- [EvaluationResultContent](#)
- [EvaluationTarget](#)
- [Evaluator](#)
- [EvaluatorMetric](#)
- [EvaluatorStatistics](#)
- [EvaluatorSummary](#)
- [Event](#)
- [EventMetadataFilterExpression](#)
- [ExecutionSummaryAffectedSession](#)
- [ExecutionSummaryCluster](#)
- [ExecutionSummaryClusteringResultContent](#)
- [ExternalProxy](#)
- [ExtractionJob](#)
- [ExtractionJobFilterInput](#)
- [ExtractionJobMessages](#)
- [ExtractionJobMetadata](#)
- [FailureAnalysisResultContent](#)
- [FailureCategoryCluster](#)
- [FailureSpanDetail](#)

- [FailureSubCategoryCluster](#)
- [FilterInput](#)
- [FilterValue](#)
- [GatewayFilter](#)
- [GroundTruthSource](#)
- [GroundTruthTurn](#)
- [GroundTruthTurnInput](#)
- [HarnessAgentCoreBrowserConfig](#)
- [HarnessAgentCoreCodeInterpreterConfig](#)
- [HarnessAgentCoreGatewayConfig](#)
- [HarnessBedrockModelConfig](#)
- [HarnessContentBlock](#)
- [HarnessContentBlockDelta](#)
- [HarnessContentBlockDeltaEvent](#)
- [HarnessContentBlockStart](#)
- [HarnessContentBlockStartEvent](#)
- [HarnessContentBlockStopEvent](#)
- [HarnessGatewayOutboundAuth](#)
- [HarnessGeminiModelConfig](#)
- [HarnessInlineFunctionConfig](#)
- [HarnessLiteLlmModelConfig](#)
- [HarnessMessage](#)
- [HarnessMessageStartEvent](#)
- [HarnessMessageStopEvent](#)
- [HarnessMetadataEvent](#)
- [HarnessModelConfiguration](#)
- [HarnessOpenAiModelConfig](#)
- [HarnessReasoningContentBlock](#)
- [HarnessReasoningContentBlockDelta](#)
- [HarnessReasoningTextBlock](#)

- [HarnessRemoteMcpConfig](#)
- [HarnessSkill](#)
- [HarnessSkillAwsSkillsSource](#)
- [HarnessSkillGitAuth](#)
- [HarnessSkillGitSource](#)
- [HarnessSkillS3Source](#)
- [HarnessStreamMetrics](#)
- [HarnessSystemContentBlock](#)
- [HarnessTokenUsage](#)
- [HarnessTool](#)
- [HarnessToolConfiguration](#)
- [HarnessToolResultBlock](#)
- [HarnessToolResultBlockDelta](#)
- [HarnessToolResultBlockStart](#)
- [HarnessToolResultContentBlock](#)
- [HarnessToolUseBlock](#)
- [HarnessToolUseBlockDelta](#)
- [HarnessToolUseBlockStart](#)
- [InlineGroundTruth](#)
- [InputContentBlock](#)
- [Insight](#)
- [InsightsFailureSignal](#)
- [InvokeAgentRuntimeCommandRequestBody](#)
- [InvokeAgentRuntimeCommandStreamOutput](#)
- [InvokeHarnessStreamOutput](#)
- [KeyPressArguments](#)
- [KeyPressResult](#)
- [KeyShortcutArguments](#)
- [KeyShortcutResult](#)
- [KeyTypeArguments](#)

- [KeyTypeResult](#)
- [LeftExpression](#)
- [LinkedAccount](#)
- [LinkedAccountDeveloperJwt](#)
- [LinkedAccountEmail](#)
- [LinkedAccountOAuth2](#)
- [LinkedAccountSms](#)
- [LiveViewStream](#)
- [McpDescriptor](#)
- [MemoryContent](#)
- [MemoryMetadataFilterExpression](#)
- [MemoryRecord](#)
- [MemoryRecordCreateInput](#)
- [MemoryRecordDeleteInput](#)
- [MemoryRecordLeftExpression](#)
- [MemoryRecordMetadataValue](#)
- [MemoryRecordOutput](#)
- [MemoryRecordRightExpression](#)
- [MemoryRecordSummary](#)
- [MemoryRecordUpdateInput](#)
- [MessageMetadata](#)
- [MetadataValue](#)
- [MouseClickedArguments](#)
- [MouseClickedResult](#)
- [MouseDownArguments](#)
- [MouseDownResult](#)
- [MouseMoveArguments](#)
- [MouseMoveResult](#)
- [MouseScrollArguments](#)
- [MouseScrollResult](#)

- [OAuth2Authentication](#)
- [OAuthCredentialProvider](#)
- [OnlineEvaluationConfigSource](#)
- [OutputConfig](#)
- [PayloadType](#)
- [PaymentInput](#)
- [PaymentInstrument](#)
- [PaymentInstrumentDetails](#)
- [PaymentInstrumentSummary](#)
- [PaymentOutput](#)
- [PaymentSession](#)
- [PaymentSessionSummary](#)
- [PaymentTokenRequestInput](#)
- [PaymentTokenResponseOutput](#)
- [PerVariantOnlineEvaluationConfig](#)
- [Proxy](#)
- [ProxyBypass](#)
- [ProxyConfiguration](#)
- [ProxyCredentials](#)
- [RecommendationConfig](#)
- [RecommendationEvaluationConfig](#)
- [RecommendationEvaluatorReference](#)
- [RecommendationResult](#)
- [RecommendationResultConfigurationBundle](#)
- [RecommendationSummary](#)
- [RegistryRecordSummary](#)
- [ResourceContent](#)
- [ResourceLocation](#)
- [ResponseChunk](#)
- [RightExpression](#)

- [RootCauseCluster](#)
- [S3Location](#)
- [ScreenshotArguments](#)
- [ScreenshotResult](#)
- [SearchCriteria](#)
- [SecretsManagerLocation](#)
- [ServerDefinition](#)
- [SessionFilter](#)
- [SessionFilterConfig](#)
- [SessionLimits](#)
- [SessionMetadataShape](#)
- [SessionSummary](#)
- [SkillDefinition](#)
- [SkillMdDefinition](#)
- [SpanContext](#)
- [StreamUpdate](#)
- [StripePrivyTokenRequestInput](#)
- [StripePrivyTokenResponseOutput](#)
- [SystemPromptConfig](#)
- [SystemPromptConfigurationBundle](#)
- [SystemPromptRecommendationConfig](#)
- [SystemPromptRecommendationResult](#)
- [TargetRef](#)
- [TokenBalance](#)
- [TokenUsage](#)
- [ToolArguments](#)
- [ToolDescriptionConfig](#)
- [ToolDescriptionConfigurationBundle](#)
- [ToolDescriptionInput](#)
- [ToolDescriptionOutput](#)

- [ToolDescriptionRecommendationConfig](#)
- [ToolDescriptionRecommendationResult](#)
- [ToolDescriptionSource](#)
- [ToolDescriptionTextInput](#)
- [ToolResultStructuredContent](#)
- [ToolsDefinition](#)
- [UserIdentifier](#)
- [UserIntentAffectedSession](#)
- [UserIntentCluster](#)
- [UserIntentClusteringResultContent](#)
- [ValidationExceptionField](#)
- [Variant](#)
- [VariantConfiguration](#)
- [VariantResult](#)
- [ViewPort](#)

A2aDescriptor

The A2A (Agent-to-Agent) descriptor configuration for a registry record.

Contents

agentCard

The agent card definition that describes the agent's capabilities and interface.

Type: [AgentCardDefinition](#) object

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ABTestEvaluationConfig

The evaluation configuration for an A/B test, specifying which online evaluation configurations to use for measuring variant performance.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

onlineEvaluationConfigArn

The Amazon Resource Name (ARN) of a single online evaluation configuration to use for both variants.

Type: String

Pattern: `arn:aws[a-zA-Z-]*:bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:online-evaluation-config\[a-zA-Z][a-zA-Z0-9-_\]{0,99}-[a-zA-Z0-9]{10}`

Required: No

perVariantOnlineEvaluationConfig

Per-variant online evaluation configurations, allowing different evaluation settings for each variant.

Type: Array of [PerVariantOnlineEvaluationConfig](#) objects

Array Members: Fixed number of 2 items.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ABTestResults

The statistical results of an A/B test.

Contents

evaluatorMetrics

The per-evaluator metrics comparing control and treatment variants.

Type: Array of [EvaluatorMetric](#) objects

Required: Yes

analysisTimestamp

The timestamp when the analysis was performed.

Type: Timestamp

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ABTestSummary

Summary information about an A/B test.

Contents

abTestArn

The Amazon Resource Name (ARN) of the A/B test.

Type: String

Pattern: `arn:aws[a-zA-Z-]*:bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:ab-test/[a-zA-Z][a-zA-Z0-9-_{0,99}]-[a-zA-Z0-9]{10}`

Required: Yes

abTestId

The unique identifier of the A/B test.

Type: String

Pattern: `[a-zA-Z][a-zA-Z0-9-_{0,99}]-[a-zA-Z0-9]{10}`

Required: Yes

createdAt

The timestamp when the A/B test was created.

Type: Timestamp

Required: Yes

executionStatus

The execution status of the A/B test.

Type: String

Valid Values: PAUSED | RUNNING | STOPPED | NOT_STARTED

Required: Yes

name

The name of the A/B test.

Type: String

Pattern: `[a-zA-Z][a-zA-Z0-9_]{0,47}`

Required: Yes

status

The current status of the A/B test.

Type: String

Valid Values: `CREATING | ACTIVE | CREATE_FAILED | UPDATING | UPDATE_FAILED | DELETING | DELETE_FAILED | FAILED`

Required: Yes

updatedAt

The timestamp when the A/B test was last updated.

Type: Timestamp

Required: Yes

description

The description of the A/B test.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 200.

Required: No

gatewayArn

The Amazon Resource Name (ARN) of the gateway used for traffic splitting.

Type: String

Pattern: `arn:aws(|-cn|-us-gov):bedrock-agentcore:[a-z0-9-]{1,20}::[0-9]{12}:gateway/([0-9a-z][-]?){1,48}-[a-z0-9]{10}`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ActorSummary

Contains summary information about an actor in an AgentCore Memory resource.

Contents

actorId

The unique identifier of the actor.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[a-zA-Z0-9][a-zA-Z0-9-_/]*(?::[a-zA-Z0-9-_/]+)*[a-zA-Z0-9-_/]*`

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

AffectedSession

A session affected by a detected failure pattern, including root cause details.

Contents

explanation

An explanation of how the failure manifested in this session.

Type: String

Required: Yes

failureSpans

The list of spans where failures were detected in this session.

Type: Array of [FailureSpanDetail](#) objects

Array Members: Minimum number of 0 items.

Required: Yes

fixType

The type of fix recommended for this failure.

Type: String

Required: Yes

recommendation

The specific fix recommendation for this session.

Type: String

Required: Yes

sessionId

The unique identifier of the affected session.

Type: String

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

AgentCardDefinition

The agent card definition for A2A descriptors, including the schema version and inline content that describes the agent's capabilities.

Contents

inlineContent

The inline content of the agent card definition.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 409600.

Required: No

schemaVersion

The schema version of the agent card definition.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

AgentSkillsDescriptor

The agent skills descriptor configuration for a registry record.

Contents

skillMd

The skill description in markdown format.

Type: [SkillMdDefinition](#) object

Required: Yes

skillDefinition

The structured skill definition with a schema version and content.

Type: [SkillDefinition](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

AgentTracesConfig

The configuration specifying where to read agent traces from for recommendation analysis.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

batchEvaluation

Use a completed batch evaluation as the source of agent traces.

Type: [BatchEvaluationTraceConfig](#) object

Required: No

cloudwatchLogs

Agent traces read from CloudWatch Logs.

Type: [CloudWatchLogsTraceConfig](#) object

Required: No

sessionSpans

Agent traces provided as inline session spans in OpenTelemetry format.

Type: Array of JSON values

Array Members: Minimum number of 1 item. Maximum number of 1000 items.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Amount

Represents a monetary amount with a currency.

Contents

currency

The currency code for the amount.

Type: String

Valid Values: USD

Required: Yes

value

The numeric value of the amount.

Type: String

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

AutomationStream

The configuration for a stream that enables programmatic control of a browser session in Amazon Bedrock AgentCore. This stream provides a bidirectional communication channel for sending commands to the browser and receiving responses, allowing agents to automate web interactions such as navigation, form filling, and element clicking.

Contents

streamEndpoint

The endpoint URL for the automation stream. This URL is used to establish a WebSocket connection to the stream for sending commands and receiving responses.

Type: String

Length Constraints: Minimum length of 10. Maximum length of 512.

Required: Yes

streamStatus

The current status of the automation stream. This indicates whether the stream is available for use. Possible values include ACTIVE, CONNECTING, and DISCONNECTED.

Type: String

Valid Values: ENABLED | DISABLED

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

AutomationStreamUpdate

Contains information about an update to an automation stream.

Contents

streamStatus

The status of the automation stream.

Type: String

Valid Values: ENABLED | DISABLED

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

AvailableLimits

The available spending limits for a payment session.

Contents

availableSpendAmount

The remaining available amount that can be spent.

Type: [Amount](#) object

Required: No

updatedAt

The timestamp when the available limits were last updated.

Type: Timestamp

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

BasicAuth

Configuration for HTTP Basic Authentication using credentials stored in AWS Secrets Manager. The secret must contain a JSON object with username and password string fields. Username allows alphanumeric characters and @._+ = - symbols (pattern: `^[a-zA-Z0-9@._+=\ -]+$`). Password allows alphanumeric characters and @._+ = - !#\$%&* symbols (pattern: `^[a-zA-Z0-9@._+=\ - !#$%&*]+$`). Both fields have a maximum length of 256 characters.

Contents

secretArn

The Amazon Resource Name (ARN) of the AWS Secrets Manager secret containing proxy credentials. The secret must be a JSON object with username and password string fields that meet validation requirements. The caller must have `secretsmanager:GetSecretValue` permission for this ARN. Example secret format: `{"username": "proxy_user", "password": "secure_password"}`

Type: String

Pattern: `arn:aws(-[a-z-]+)?:secretsmanager:[a-z0-9-]+:[0-9]{12}:secret:[a-zA-Z0-9/_+=.@-]+`

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

BatchEvaluationSummary

Summary representation for list responses.

Contents

batchEvaluationArn

The Amazon Resource Name (ARN) of the batch evaluation.

Type: String

Required: Yes

batchEvaluationId

The unique identifier of the batch evaluation.

Type: String

Pattern: `[a-zA-Z][a-zA-Z0-9-_{0,99}-[a-zA-Z0-9]{10}`

Required: Yes

batchEvaluationName

The name of the batch evaluation.

Type: String

Pattern: `[a-zA-Z][a-zA-Z0-9_]{0,47}`

Required: Yes

createdAt

The timestamp when the batch evaluation was created.

Type: Timestamp

Required: Yes

status

The current status of the batch evaluation.

Type: String

Valid Values: PENDING | IN_PROGRESS | COMPLETED | COMPLETED_WITH_ERRORS | FAILED | STOPPING | STOPPED | DELETING

Required: Yes

description

The description of the batch evaluation.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 200.

Required: No

errorDetails

The error details if the batch evaluation encountered failures.

Type: Array of strings

Array Members: Minimum number of 0 items. Maximum number of 1 item.

Length Constraints: Minimum length of 0. Maximum length of 1000.

Required: No

evaluationResults

The aggregated evaluation results.

Type: [EvaluationJobResults](#) object

Required: No

evaluators

The list of evaluators applied during the batch evaluation.

Type: Array of [Evaluator](#) objects

Required: No

insights

The list of insight analyses applied during the batch evaluation.

Type: Array of [Insight](#) objects

Array Members: Minimum number of 0 items. Maximum number of 10 items.

Required: No

kmsKeyArn

The ARN of the AWS KMS key used to encrypt evaluation data.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: `arn:aws(|-cn|-us-gov):kms:[a-zA-Z0-9-]*:[0-9]{12}:key/[a-zA-Z0-9-]{36}`

Required: No

updatedAt

The timestamp when the batch evaluation was last updated.

Type: Timestamp

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

BatchEvaluationTraceConfig

Configuration for using a batch evaluation as the source of agent traces for recommendations.

Contents

batchEvaluationArn

The ARN of the completed batch evaluation to use as the trace source.

Type: String

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Branch

Contains information about a branch in an AgentCore Memory resource. Branches allow for organizing events into different conversation threads or paths.

Contents

name

The name of the branch.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 100.

Pattern: `[a-zA-Z0-9][a-zA-Z0-9-_*]`

Required: Yes

rootEventId

The identifier of the root event for this branch.

Type: String

Pattern: `[0-9]+#[a-fA-F0-9]+`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

BranchFilter

Contains filter criteria for branches when listing events.

Contents

name

The name of the branch to filter by.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 100.

Pattern: `[a-zA-Z0-9][a-zA-Z0-9-_*]`

Required: Yes

includeParentBranches

Specifies whether to include parent branches in the results. Set to true to include parent branches, or false to exclude them.

Type: Boolean

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

BrowserAction

The browser action to perform. Exactly one member must be set per request.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

keyPress

Press a key one or more times.

Type: [KeyPressArguments](#) object

Required: No

keyShortcut

Press a key combination.

Type: [KeyShortcutArguments](#) object

Required: No

keyType

Type a string of text.

Type: [KeyTypeArguments](#) object

Required: No

mouseClick

Click at the specified coordinates.

Type: [MouseClickedArguments](#) object

Required: No

mouseDrag

Drag from a start position to an end position.

Type: [MouseDownArguments](#) object

Required: No

mouseMove

Move the cursor to the specified coordinates.

Type: [MouseMoveArguments](#) object

Required: No

mouseScroll

Scroll at the specified position.

Type: [MouseScrollArguments](#) object

Required: No

screenshot

Capture a full-screen screenshot.

Type: [ScreenshotArguments](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

BrowserActionResult

The result of a browser action execution. Exactly one member is set, matching the action that was performed.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

keyPress

The result of a key press action.

Type: [KeyPressResult](#) object

Required: No

keyShortcut

The result of a key shortcut action.

Type: [KeyShortcutResult](#) object

Required: No

keyType

The result of a key type action.

Type: [KeyTypeResult](#) object

Required: No

mouseClick

The result of a mouse click action.

Type: [MouseClickedResult](#) object

Required: No

mouseDrag

The result of a mouse drag action.

Type: [MouseDownResult](#) object

Required: No

mouseMove

The result of a mouse move action.

Type: [MouseMoveResult](#) object

Required: No

mouseScroll

The result of a mouse scroll action.

Type: [MouseScrollResult](#) object

Required: No

screenshot

The result of a screenshot action.

Type: [ScreenshotResult](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

BrowserEnterprisePolicy

Browser enterprise policy configuration.

Contents

location

The location of the enterprise policy file.

Type: [ResourceLocation](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: Yes

type

The enterprise policy type. See [BrowserEnterprisePolicyType](#).

Type: String

Valid Values: MANAGED | RECOMMENDED

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

BrowserExtension

Browser extension configuration.

Contents

location

The location where the browser extension files are stored. This specifies the source from which the extension will be loaded and installed.

Type: [ResourceLocation](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

BrowserProfileConfiguration

The configuration for a browser profile in Amazon Bedrock AgentCore. A browser profile contains persistent browser data such as cookies and local storage that can be saved from one browser session and reused in subsequent sessions. Browser profiles enable continuity for tasks that require authentication, maintain user preferences, or depend on previously stored browser state.

Contents

profileIdentifier

The unique identifier of the browser profile. This identifier is used to reference the profile when starting new browser sessions or saving session data to the profile.

Type: String

Pattern: `[a-zA-Z][a-zA-Z0-9_]{0,47}-[a-zA-Z0-9]{10}`

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

BrowserSessionStream

The collection of streams associated with a browser session in Amazon Bedrock AgentCore. These streams provide different ways to interact with and observe the browser session, including programmatic control and visual representation of the browser content.

Contents

automationStream

The stream that enables programmatic control of the browser. This stream allows agents to perform actions such as navigating to URLs, clicking elements, and filling forms.

Type: [AutomationStream](#) object

Required: Yes

liveViewStream

The stream that provides a visual representation of the browser content. This stream allows agents to observe the current state of the browser, including rendered web pages and visual elements.

Type: [LiveViewStream](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

BrowserSessionSummary

A condensed representation of a browser session in Amazon Bedrock AgentCore. This structure contains key information about a browser session, including identifiers, status, and timestamps, without the full details of the session configuration and streams.

Contents

browserIdentifier

The unique identifier of the browser associated with the session. This identifier specifies which browser environment is used for the session.

Type: String

Required: Yes

createdAt

The timestamp when the browser session was created. This value is in ISO 8601 format.

Type: Timestamp

Required: Yes

sessionId

The unique identifier of the browser session. This identifier is used in operations that interact with the session.

Type: String

Pattern: `[0-9a-zA-Z]{1,40}`

Required: Yes

status

The current status of the browser session. Possible values include ACTIVE, STOPPING, and STOPPED.

Type: String

Valid Values: READY | TERMINATED

Required: Yes

lastUpdatedAt

The timestamp when the browser session was last updated. This value is in ISO 8601 format.

Type: Timestamp

Required: No

name

The name of the browser session. This name helps identify and manage the session.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 100.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Certificate

A certificate to install in the browser or code interpreter session.

Contents

location

The location of the certificate.

Type: [CertificateLocation](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

CertificateLocation

The location from which to retrieve a certificate.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

secretsManager

The AWS Secrets Manager location of the certificate.

Type: [SecretsManagerLocation](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

CloudWatchFilterConfig

Filter configuration for narrowing down CloudWatch Logs sessions for evaluation.

Contents

sessionIds

A list of specific session IDs to evaluate. If specified, only these sessions are included in the evaluation.

Type: Array of strings

Array Members: Minimum number of 0 items. Maximum number of 500 items.

Required: No

timeRange

The time range filter for selecting sessions to evaluate.

Type: [SessionFilterConfig](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

CloudWatchLogsFilter

A filter for narrowing down agent traces from CloudWatch Logs based on key-value comparisons.

Contents

key

The key or field name to filter on within the agent trace data.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: `[a-zA-Z0-9._-]+`

Required: Yes

operator

The comparison operator to use for filtering.

Type: String

Valid Values: `Equals` | `NotEquals` | `GreaterThan` | `LessThan` | `GreaterThanOrEqual` | `LessThanOrEqual` | `Contains` | `NotContains`

Required: Yes

value

The value to compare against using the specified operator.

Type: [FilterValue](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

CloudWatchLogsRule

A rule configuration for filtering agent traces from CloudWatch Logs.

Contents

filters

The list of filters to apply when reading agent traces.

Type: Array of [CloudWatchLogsFilter](#) objects

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

CloudWatchLogsSource

The configuration for reading agent traces from CloudWatch Logs.

Contents

logGroupNames

The list of CloudWatch log group names to read agent traces from. Maximum of 5 log groups.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 5 items.

Required: Yes

serviceNames

The list of agent service names to filter traces within the specified log groups.

Type: Array of strings

Array Members: Fixed number of 1 item.

Required: Yes

filterConfig

Optional filter configuration to narrow down which sessions to evaluate.

Type: [CloudWatchFilterConfig](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

CloudWatchLogsTraceConfig

Configuration for reading agent traces from CloudWatch Logs for recommendation analysis.

Contents

endTime

The end time of the time range to read traces from.

Type: Timestamp

Required: Yes

logGroupArns

The list of CloudWatch log group ARNs to read agent traces from.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 5 items.

Required: Yes

serviceNames

The list of service names to filter traces within the specified log groups.

Type: Array of strings

Array Members: Fixed number of 1 item.

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: [a-zA-Z0-9._-]+

Required: Yes

startTime

The start time of the time range to read traces from.

Type: Timestamp

Required: Yes

rule

Optional rule configuration for filtering traces.

Type: [CloudWatchLogsRule](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

CloudWatchOutputConfig

CloudWatch Logs destination for batch evaluation results.

Contents

logGroupName

The name of the CloudWatch log group where evaluation results will be written.

Type: String

Required: Yes

logStreamName

The name of the CloudWatch log stream where evaluation results will be written.

Type: String

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

CodeInterpreterResult

The output produced by executing code in a code interpreter session in Amazon Bedrock AgentCore. This structure contains the results of code execution, including textual output, structured data, and error information. Agents use these results to generate responses that incorporate computation, data analysis, and visualization.

Contents

content

The textual content of the execution result. This includes standard output from the code execution, such as print statements, console output, and text representations of results.

Type: Array of [ContentBlock](#) objects

Required: Yes

isError

Indicates whether the result represents an error. If true, the content contains error messages or exception information. If false, the content contains successful execution results.

Type: Boolean

Required: No

structuredContent

The structured content of the execution result. This includes additional metadata about the execution, such as execution time, memory usage, and structured representations of output data. The format depends on the specific code interpreter and execution context.

Type: [ToolResultStructuredContent](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

CodeInterpreterSessionSummary

A condensed representation of a code interpreter session in Amazon Bedrock AgentCore. This structure contains key information about a code interpreter session, including identifiers, status, and timestamps, without the full details of the session configuration.

Contents

codeInterpreterIdentifier

The unique identifier of the code interpreter associated with the session. This identifier specifies which code interpreter environment is used for the session.

Type: String

Required: Yes

createdAt

The timestamp when the code interpreter session was created. This value is in ISO 8601 format.

Type: Timestamp

Required: Yes

sessionId

The unique identifier of the code interpreter session. This identifier is used in operations that interact with the session.

Type: String

Pattern: `[0-9a-zA-Z]{1,40}`

Required: Yes

status

The current status of the code interpreter session. Possible values include ACTIVE, STOPPING, and STOPPED.

Type: String

Valid Values: READY | TERMINATED

Required: Yes

lastUpdatedAt

The timestamp when the code interpreter session was last updated. This value is in ISO 8601 format.

Type: Timestamp

Required: No

name

The name of the code interpreter session. This name helps identify and manage the session.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 100.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

CodeInterpreterStreamOutput

Contains output from a code interpreter stream.

Contents

accessDeniedException

The exception that occurs when you do not have sufficient permissions to perform an action. Verify that your IAM policy includes the necessary permissions for the operation you are trying to perform.

Type: Exception

HTTP Status Code: 403

Required: No

conflictException

The exception that occurs when the request conflicts with the current state of the resource. This can happen when trying to modify a resource that is currently being modified by another request, or when trying to create a resource that already exists.

Type: Exception

HTTP Status Code: 409

Required: No

internalServerErrorException

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

Type: Exception

HTTP Status Code: 500

Required: No

resourceNotFoundException

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

Type: Exception

HTTP Status Code: 404

Required: No

result

The output produced by executing code in a code interpreter session in Amazon Bedrock AgentCore. This structure contains the results of code execution, including textual output, structured data, and error information. Agents use these results to generate responses that incorporate computation, data analysis, and visualization.

Type: [CodeInterpreterResult](#) object

Required: No

serviceQuotaExceededException

The exception that occurs when the request would cause a service quota to be exceeded. Review your service quotas and either reduce your request rate or request a quota increase.

Type: Exception

HTTP Status Code: 402

Required: No

throttlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

Type: Exception

HTTP Status Code: 429

Required: No

validationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

Type: Exception

HTTP Status Code: 400

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

CoinbaseCdpTokenRequestInput

Coinbase CDP token request parameters.

Contents

requestMethod

The HTTP method for the payment API request.

Type: String

Valid Values: GET | POST | PUT | DELETE | PATCH

Required: Yes

requestPath

The path of the payment API request.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: `/[a-zA-Z0-9/_\-\.\~%?=&]+`

Required: Yes

includeWalletAuthToken

Set to true for wallet write operations (requires walletSecret configured).

Type: Boolean

Required: No

requestBody

Request body JSON — used to generate wallet auth JWT.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 16384.

Pattern: `[\u0009\u000A\u000D\u0020-\u007E]+`

Required: No

requestHost

The host for the payment API request. Defaults to "api.cdp.coinbase.com".

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: [a-zA-Z0-9\-\.\]+

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

CoinbaseCdpTokenResponseOutput

Coinbase CDP token response.

Contents

bearerToken

Bearer Token for Authorization header.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 8192.

Required: Yes

walletAuthToken

Wallet Auth Token for X-Wallet-Auth header.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 8192.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ConfidenceInterval

A confidence interval for a statistical measurement.

Contents

lower

The lower bound of the confidence interval.

Type: Double

Required: No

upper

The upper bound of the confidence interval.

Type: Double

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ConfigurationBundleRef

A reference to a specific version of a configuration bundle.

Contents

bundleArn

The Amazon Resource Name (ARN) of the configuration bundle.

Type: String

Pattern: `arn:aws[a-zA-Z-]*:bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:configuration-bundle/[a-zA-Z][a-zA-Z0-9-_{0,99}]-[a-zA-Z0-9]{10}`

Required: Yes

bundleVersion

The version of the configuration bundle.

Type: String

Pattern: `[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}`

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ConfigurationBundleToolEntry

Maps a tool name to its JSON path within a configuration bundle.

Contents

toolDescriptionJsonPath

The JSON path within the configuration bundle's components that contains the tool description.

Type: String

Required: Yes

toolName

The name of the tool.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: `[a-zA-Z0-9_-\.\]+`

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Content

Contains the content of a memory item.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

text

The text content of the memory item.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 100000.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ContentBlock

A block of content in a response.

Contents

type

The type of content in the block.

Type: String

Valid Values: `text` | `image` | `resource` | `resource_link`

Required: Yes

data

The binary data content of the block.

Type: Base64-encoded binary data object

Required: No

description

The description of the content block.

Type: String

Required: No

mimeType

The MIME type of the content.

Type: String

Required: No

name

The name of the content block.

Type: String

Required: No

resource

The resource associated with the content block.

Type: [ResourceContent](#) object

Required: No

size

The size of the content in bytes.

Type: Long

Required: No

text

The text content of the block.

Type: String

Required: No

uri

The URI of the content.

Type: String

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ContentDeltaEvent

An event that contains incremental output from a command execution. This event streams standard output and standard error content as it becomes available during command execution.

Contents

stderr

The standard error content from the command execution. This field contains the incremental output written to stderr by the executing command.

Type: String

Required: No

stdout

The standard output content from the command execution. This field contains the incremental output written to stdout by the executing command.

Type: String

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ContentStartEvent

An event that signals the start of content streaming from a command execution. This event is sent when the command begins producing output.

Contents

The members of this exception structure are context-dependent.

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ContentStopEvent

An event that signals the completion of a command execution. This event contains the final status and exit code of the executed command.

Contents

exitCode

The exit code returned by the executed command. An exit code of 0 indicates successful execution, -1 indicates a platform error, and values greater than 0 indicate command-specific errors.

Type: Integer

Required: Yes

status

The final status of the command execution. Valid values are COMPLETED for successful completion or TIMED_OUT if the command exceeded the specified timeout.

Type: String

Valid Values: COMPLETED | TIMED_OUT

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Context

The contextual information associated with an evaluation, including span context details that identify the specific traces and sessions being evaluated within the agent's execution flow.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

spanContext

The span context information that uniquely identifies the trace and span being evaluated, including session ID, trace ID, and span ID for precise targeting within the agent's execution flow.

Type: [SpanContext](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ControlStats

Statistics for the control variant in an A/B test.

Contents

mean

The mean evaluation score for the control variant.

Type: Double

Required: Yes

sampleSize

The number of sessions evaluated for the control variant.

Type: Integer

Required: Yes

variantName

The name of the control variant.

Type: String

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Conversational

Contains conversational content for an event payload.

Contents

content

The content of the conversation message.

Type: [Content](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: Yes

role

The role of the participant in the conversation (for example, "user" or "assistant").

Type: String

Valid Values: ASSISTANT | USER | TOOL | OTHER

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

CryptoX402PaymentInput

The input for a crypto X402 payment.

Contents

payload

The X402 payment payload.

Type: JSON value

Required: Yes

version

The version of the X402 protocol.

Type: String

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

CryptoX402PaymentOutput

The output from a crypto X402 payment.

Contents

payload

The X402 payment response payload.

Type: JSON value

Required: Yes

version

The version of the X402 protocol.

Type: String

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

CustomDescriptor

A custom descriptor configuration for a registry record.

Contents

inlineContent

The inline content of the custom descriptor.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 409600.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

DataSourceConfig

Configuration for the data source used in evaluation.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

cloudWatchLogs

Configuration for pulling agent session traces from CloudWatch Logs.

Type: [CloudWatchLogsSource](#) object

Required: No

onlineEvaluationConfigSource

A reference to an existing online evaluation configuration to use as the data source for batch evaluation.

Type: [OnlineEvaluationConfigSource](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Descriptors

Contains the descriptor configuration for a registry record. Only the field that matches the record's `descriptorType` is populated.

Contents

a2a

The A2A (Agent-to-Agent) descriptor configuration. Populated when the record's `descriptorType` is A2A.

Type: [A2aDescriptor](#) object

Required: No

agentSkills

The agent skills descriptor configuration. Populated when the record's `descriptorType` is AGENT_SKILLS.

Type: [AgentSkillsDescriptor](#) object

Required: No

custom

The custom descriptor configuration. Populated when the record's `descriptorType` is CUSTOM.

Type: [CustomDescriptor](#) object

Required: No

mcp

The MCP (Model Context Protocol) descriptor configuration. Populated when the record's `descriptorType` is MCP.

Type: [McpDescriptor](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

EmbeddedCryptoWallet

Embedded crypto wallet instrument details.

Contents

linkedAccounts

List of linked accounts linked to this wallet. Each represents a way the end user can authenticate to this wallet.

Type: Array of [LinkedAccount](#) objects

Array Members: Minimum number of 0 items. Maximum number of 1 item.

Required: Yes

network

The blockchain network for this embedded crypto wallet. Supported networks: ETHEREUM, SOLANA.

Type: String

Valid Values: ETHEREUM | SOLANA

Required: Yes

redirectUrl

URL for the end user to complete a provider-specific action such as wallet linking or onboarding.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Required: No

walletAddress

The wallet address on the specified blockchain network.

Type: String

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

EvaluationContent

A content block for ground truth data in evaluation reference inputs. Supports text content for expected responses and assertions.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

text

The text content of the ground truth data. Used for expected response text and assertion statements.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 100000.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

EvaluationExpectedTrajectory

The expected tool call trajectory for trajectory-based evaluation.

Contents

toolNames

The list of tool names representing the expected tool call sequence.

Type: Array of strings

Array Members: Minimum number of 0 items. Maximum number of 1000 items.

Length Constraints: Minimum length of 1. Maximum length of 500.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

EvaluationInput

The input data structure containing agent session spans in OpenTelemetry format. Supports traces from frameworks like Strands (AgentCore Runtime) and LangGraph with OpenInference instrumentation for comprehensive evaluation.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

sessionSpans

The collection of spans representing agent execution traces within a session. Each span contains detailed information about tool calls, model interactions, and other agent activities that can be evaluated for quality and performance.

Type: Array of JSON values

Array Members: Minimum number of 1 item. Maximum number of 1000 items.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

EvaluationJobResults

Aggregated results from a batch evaluation, including session completion counts and evaluator score summaries.

Contents

evaluatorSummaries

A list of per-evaluator summary statistics.

Type: Array of [EvaluatorSummary](#) objects

Required: No

numberOfSessionsCompleted

The number of sessions that have been successfully evaluated.

Type: Integer

Required: No

numberOfSessionsFailed

The number of sessions that failed evaluation.

Type: Integer

Required: No

numberOfSessionsIgnored

The number of sessions that were ignored during evaluation.

Type: Integer

Required: No

numberOfSessionsInProgress

The number of sessions currently being evaluated.

Type: Integer

Required: No

totalNumberOfSessions

The total number of sessions included in the batch evaluation.

Type: Integer

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

EvaluationMetadata

Metadata for the evaluation, including session-specific ground truth data.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

sessionMetadata

A list of session metadata entries containing ground truth data and test scenario identifiers for specific sessions.

Type: Array of [SessionMetadataShape](#) objects

Array Members: Minimum number of 0 items. Maximum number of 500 items.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

EvaluationReferenceInput

A reference input containing ground truth data for evaluation, scoped to a specific context level (session or trace) through its span context.

Contents

context

The span context that identifies which session or trace this reference input applies to, used for correlating ground truth with agent output.

Type: [Context](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: Yes

assertions

A list of assertion statements for session-level evaluation. Each assertion describes an expected behavior or outcome the agent should demonstrate during the session.

Type: Array of [EvaluationContent](#) objects

Array Members: Minimum number of 1 item. Maximum number of 100 items.

Required: No

expectedResponse

The expected response for trace-level evaluation. Built-in evaluators that support this field compare the agent's actual response against this value for assessment. Custom evaluators can access it through the `{expected_response}` placeholder in their instructions.

Type: [EvaluationContent](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: No

expectedTrajectory

The expected tool call sequence for session-level trajectory evaluation. Contains a list of tool names representing the tools the agent is expected to invoke.

Type: [EvaluationExpectedTrajectory](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

EvaluationResultContent

The comprehensive result of an evaluation containing the score, explanation, evaluator metadata, and execution details. Provides both quantitative ratings and qualitative insights about agent performance.

Contents

context

The contextual information associated with this evaluation result, including span context details that identify the specific traces and sessions that were evaluated.

Type: [Context](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: Yes

evaluatorArn

The Amazon Resource Name (ARN) of the evaluator used to generate this result. For custom evaluators, this is the full ARN; for built-in evaluators, this follows the pattern `Builtin.{EvaluatorName}`.

Type: String

Pattern: `arn:aws[a-zA-Z-]*:bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:evaluator\|[a-zA-Z][a-zA-Z0-9-]{0,99}-[a-zA-Z0-9]{10}$|^arn:aws[a-zA-Z-]*:bedrock-agentcore::evaluator/Builtin.[a-zA-Z0-9_-]+`

Required: Yes

evaluatorId

The unique identifier of the evaluator that produced this result. This matches the `evaluatorId` provided in the evaluation request and can be used to identify which evaluator generated specific results.

Type: String

Pattern: `(Builtin.[a-zA-Z0-9_-]+|[a-zA-Z][a-zA-Z0-9-]{0,99}-[a-zA-Z0-9]{10})`

Required: Yes

evaluatorName

The human-readable name of the evaluator used for this evaluation. For built-in evaluators, this is the descriptive name (e.g., "Helpfulness", "Correctness"); for custom evaluators, this is the user-defined name.

Type: String

Pattern: `(Builtin.[a-zA-Z0-9_-]+|[a-zA-Z][a-zA-Z0-9_]{0,47})`

Required: Yes

errorCode

The error code indicating the type of failure that occurred during evaluation. Used to programmatically identify and handle different categories of evaluation errors.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1024.

Required: No

errorMessage

The error message describing what went wrong if the evaluation failed. Provides detailed information about evaluation failures to help diagnose and resolve issues with evaluator configuration or input data.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Required: No

explanation

The detailed explanation provided by the evaluator describing the reasoning behind the assigned score. This qualitative feedback helps understand why specific ratings were given and provides actionable insights for improvement.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Required: No

ignoredReferenceInputFields

The list of reference input field names that were provided but not used by the evaluator. Helps identify which ground truth data was not consumed during evaluation.

Type: Array of strings

Array Members: Minimum number of 0 items. Maximum number of 100 items.

Length Constraints: Minimum length of 1. Maximum length of 1000.

Required: No

label

The categorical label assigned by the evaluator when using a categorical rating scale. This provides a human-readable description of the evaluation result (e.g., "Excellent", "Good", "Poor") corresponding to the numerical value. For numerical scales, this field is optional and provides a natural language explanation of what the value means (e.g., value 0.5 = "Somewhat Helpful").

Type: String

Required: No

tokenUsage

The token consumption statistics for this evaluation, including input tokens, output tokens, and total tokens used by the underlying language model during the evaluation process.

Type: [TokenUsage](#) object

Required: No

value

The numerical score assigned by the evaluator according to its configured rating scale. For numerical scales, this is a decimal value within the defined range. This field is not allowed for categorical scales.

Type: Double

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

EvaluationTarget

The specification of which trace or span IDs to evaluate within the provided input data. Allows precise targeting of evaluation at different levels: tool calls, traces, or sessions.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

spanIds

The list of specific span IDs to evaluate within the provided traces. Used to target evaluation at individual tool calls or specific operations within the agent's execution flow.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 10 items.

Length Constraints: Fixed length of 16.

Required: No

tracIds

The list of trace IDs to evaluate, representing complete request-response interactions. Used to evaluate entire conversation turns or specific agent interactions within a session.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 10 items.

Length Constraints: Fixed length of 32.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Evaluator

An evaluator to run against sessions during batch evaluation.

Contents

`evaluatorId`

The unique identifier of the evaluator. Can reference built-in evaluators (e.g., `Builtin.Helpfulness`) or custom evaluators.

Type: String

Pattern: `(Builtin.[a-zA-Z0-9_-]+|[a-zA-Z][a-zA-Z0-9_-]{0,99})-[a-zA-Z0-9]{10}`

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

EvaluatorMetric

Statistical metrics for a single evaluator comparing control and treatment variants.

Contents

controlStats

The statistics for the control variant.

Type: [ControlStats](#) object

Required: Yes

evaluatorArn

The Amazon Resource Name (ARN) of the evaluator.

Type: String

Required: Yes

variantResults

The results for each treatment variant compared against the control.

Type: Array of [VariantResult](#) objects

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

EvaluatorStatistics

Aggregated statistics for an evaluator.

Contents

averageScore

The average score across all evaluated sessions for this evaluator.

Type: Double

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

EvaluatorSummary

Summary statistics for a single evaluator within a batch evaluation.

Contents

evaluatorId

The unique identifier of the evaluator.

Type: String

Required: No

statistics

The aggregated statistics for this evaluator.

Type: [EvaluatorStatistics](#) object

Required: No

totalEvaluated

The total number of sessions evaluated by this evaluator.

Type: Integer

Required: No

totalFailed

The total number of sessions that failed evaluation by this evaluator.

Type: Integer

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)

- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Event

Contains information about an event in an AgentCore Memory resource.

Contents

actorId

The identifier of the actor associated with the event.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[a-zA-Z0-9][a-zA-Z0-9-_/]*(?::[a-zA-Z0-9-_/]+)*[a-zA-Z0-9-_/]*`

Required: Yes

eventId

The unique identifier of the event.

Type: String

Pattern: `[0-9]+#[a-zA-F0-9]+`

Required: Yes

eventTimestamp

The timestamp when the event occurred.

Type: Timestamp

Required: Yes

memoryId

The identifier of the AgentCore Memory resource containing the event.

Type: String

Length Constraints: Minimum length of 12.

Pattern: `(arn:(aws|aws-cn|aws-us-gov):bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:memory/)?[a-zA-Z][a-zA-Z0-9-_]{0,99}-[a-zA-Z0-9]{10}`

Required: Yes

payload

The content payload of the event.

Type: Array of [PayloadType](#) objects

Array Members: Minimum number of 0 items. Maximum number of 100 items.

Required: Yes

sessionId

The identifier of the session containing the event.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 100.

Pattern: `[a-zA-Z0-9][a-zA-Z0-9-_*]`

Required: Yes

branch

The branch information for the event.

Type: [Branch](#) object

Required: No

metadata

Metadata associated with an event.

Type: String to [MetadataValue](#) object map

Map Entries: Minimum number of 0 items. Maximum number of 15 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Key Pattern: `[a-zA-Z0-9\s._:/=+@-]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

EventMetadataFilterExpression

Filter expression for retrieving events based on metadata associated with an event.

Contents

left

Left operand of the event metadata filter expression.

Type: [LeftExpression](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: Yes

operator

Operator applied to the event metadata filter expression.

Type: String

Valid Values: EQUALS_TO | EXISTS | NOT_EXISTS

Required: Yes

right

Right operand of the event metadata filter expression.

Type: [RightExpression](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)

- [AWS SDK for Ruby V3](#)

ExecutionSummaryAffectedSession

A session associated with an execution summary cluster.

Contents

approachTaken

The approach taken by the agent during this session.

Type: String

Required: Yes

finalOutcome

The final outcome of the session.

Type: String

Required: Yes

sessionId

The unique identifier of the session.

Type: String

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ExecutionSummaryCluster

A cluster of similar execution patterns identified across sessions.

Contents

affectedSessionCount

The number of sessions with this execution pattern.

Type: Integer

Required: Yes

affectedSessions

The list of sessions with this execution pattern.

Type: Array of [ExecutionSummaryAffectedSession](#) objects

Array Members: Minimum number of 0 items.

Required: Yes

clusterId

The unique identifier of the execution summary cluster.

Type: Integer

Required: Yes

description

A description of the execution pattern.

Type: String

Required: Yes

name

The name of the execution pattern cluster.

Type: String

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ExecutionSummaryClusteringResultContent

The execution summary clustering result containing grouped execution patterns identified across evaluated sessions.

Contents

executionSummaries

The list of execution summary clusters identified across analyzed sessions.

Type: Array of [ExecutionSummaryCluster](#) objects

Array Members: Minimum number of 0 items.

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ExternalProxy

Configuration for a customer-managed external proxy server. Includes server location, optional domain-based routing patterns, and authentication credentials.

Contents

port

The port number of the proxy server. Valid range: 1-65535.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 65535.

Required: Yes

server

The hostname of the proxy server. Must be a valid DNS hostname (maximum 253 characters).

Type: String

Length Constraints: Minimum length of 1. Maximum length of 253.

Pattern: `[a-zA-Z0-9]([a-zA-Z0-9\-\]{0,61}[a-zA-Z0-9])?(\.[a-zA-Z0-9]([a-zA-Z0-9\-\]{0,61}[a-zA-Z0-9])?)*`

Required: Yes

credentials

Optional authentication credentials for the proxy server. If omitted, the proxy is accessed without authentication (useful for IP-allowlisted proxies).

Type: [ProxyCredentials](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: No

domainPatterns

Optional array of domain patterns that should route through this specific proxy. Supports `.example.com` for subdomain matching (matches any subdomain of example.com) or

example.com for exact domain matching. If omitted, this proxy acts as a catch-all for domains not matched by other proxies. Maximum 100 patterns per proxy, each up to 253 characters.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 100 items.

Length Constraints: Minimum length of 1. Maximum length of 253.

Pattern: $(\backslash.)?[a-zA-Z0-9]([a-zA-Z0-9\-\]{0,61}[a-zA-Z0-9])?(\backslash.[a-zA-Z0-9]([a-zA-Z0-9\-\]{0,61}[a-zA-Z0-9]))?^*$

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ExtractionJob

Represents the metadata of a memory extraction job such as the message identifiers that compose this job.

Contents

jobId

The unique identifier of the extraction job.

Type: String

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ExtractionJobFilterInput

Filters for querying memory extraction jobs based on various criteria.

Contents

actorId

The identifier of the actor. If specified, only extraction jobs with this actor ID are returned.

Type: String

Required: No

sessionId

The unique identifier of the session. If specified, only extraction jobs with this session ID are returned.

Type: String

Required: No

status

The status of the extraction job. If specified, only extraction jobs with this status are returned.

Type: String

Valid Values: FAILED

Required: No

strategyId

The memory strategy identifier to filter extraction jobs by. If specified, only extraction jobs with this strategy ID are returned.

Type: String

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ExtractionJobMessages

The list of messages that compose this extraction job.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

messagesList

The list of messages that compose this extraction job.

Type: Array of [MessageMetadata](#) objects

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ExtractionJobMetadata

Metadata information associated with this extraction job.

Contents

jobID

The unique identifier for the extraction job.

Type: String

Required: Yes

messages

The messages associated with the extraction job.

Type: [ExtractionJobMessages](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: Yes

actorId

The identifier of the actor for this extraction job.

Type: String

Required: No

failureReason

The cause of failure, if the job did not complete successfully.

Type: String

Required: No

sessionId

The identifier of the session for this extraction job.

Type: String

Required: No

status

The current status of the extraction job.

Type: String

Valid Values: FAILED

Required: No

strategyId

The identifier of the memory strategy for this extraction job.

Type: String

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

FailureAnalysisResultContent

The failure analysis clustering result containing categorized failure clusters with root causes and remediation recommendations.

Contents

failures

The list of failure category clusters identified across analyzed sessions.

Type: Array of [FailureCategoryCluster](#) objects

Array Members: Minimum number of 0 items.

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

FailureCategoryCluster

A top-level failure category identified by clustering similar failure patterns across sessions.

Contents

affectedSessionCount

The number of sessions affected by this failure category.

Type: Integer

Required: Yes

clusterId

The unique identifier of the failure category cluster.

Type: Integer

Required: Yes

description

A description of the failure category pattern.

Type: String

Required: Yes

name

The name of the failure category.

Type: String

Required: Yes

subCategories

The list of failure subcategories within this category.

Type: Array of [FailureSubCategoryCluster](#) objects

Array Members: Minimum number of 0 items.

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

FailureSpanDetail

Details about a specific span where a failure was detected.

Contents

signals

The failure signals detected in this span.

Type: Array of [InsightsFailureSignal](#) objects

Required: Yes

spanId

The unique identifier of the span where the failure occurred.

Type: String

Required: Yes

traceId

The trace identifier associated with the failure span.

Type: String

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

FailureSubCategoryCluster

A subcategory of failures within a top-level failure category.

Contents

affectedSessionCount

The number of sessions affected by this failure subcategory.

Type: Integer

Required: Yes

clusterId

The unique identifier of the failure subcategory cluster.

Type: Integer

Required: Yes

description

A description of the failure subcategory pattern.

Type: String

Required: Yes

name

The name of the failure subcategory.

Type: String

Required: Yes

rootCauses

The list of root cause clusters identified within this subcategory.

Type: Array of [RootCauseCluster](#) objects

Array Members: Minimum number of 0 items.

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

FilterInput

Contains filter criteria for listing events.

Contents

branch

The branch filter criteria to apply when listing events.

Type: [BranchFilter](#) object

Required: No

eventMetadata

Event metadata filter criteria to apply when retrieving events.

Type: Array of [EventMetadataFilterExpression](#) objects

Array Members: Minimum number of 1 item. Maximum number of 5 items.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

FilterValue

A value used in filter comparisons, supporting different data types.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

booleanValue

A boolean value for true/false filtering conditions.

Type: Boolean

Required: No

doubleValue

A numeric value for numerical filtering and comparisons.

Type: Double

Required: No

stringValue

A string value for text-based filtering.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 1024.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

GatewayFilter

A filter to restrict which gateway target paths are included in the A/B test.

Contents

targetPaths

A list of target path patterns to include in the A/B test.

Type: Array of strings

Array Members: Fixed number of 1 item.

Length Constraints: Minimum length of 1. Maximum length of 500.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

GroundTruthSource

Where to pull ground truth from.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

inline

Inline ground truth data provided directly in the request.

Type: [InlineGroundTruth](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

GroundTruthTurn

Ground truth data for a single conversation turn.

Contents

expectedResponse

The expected response for this conversation turn.

Type: [EvaluationContent](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: No

input

The input for this conversation turn.

Type: [GroundTruthTurnInput](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

GroundTruthTurnInput

The input for a ground truth conversation turn.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

prompt

The text prompt for this conversation turn.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 4000.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

HarnessAgentCoreBrowserConfig

Configuration for AgentCore Browser.

Contents

browserArn

If not populated, the built-in Browser ARN is used.

Type: String

Pattern: `arn:aws(-[^\:]+)?:bedrock-agentcore:[a-z0-9-]+:(aws|[0-9]{12}):browser(-custom)?/(aws\.browser\.v1|[a-zA-Z][a-zA-Z0-9_]{0,47}-[a-zA-Z0-9]{10})`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

HarnessAgentCoreCodeInterpreterConfig

Configuration for AgentCore Code Interpreter.

Contents

codeInterpreterArn

If not populated, the built-in Code Interpreter ARN is used.

Type: String

Pattern: `arn:aws(-[^\:]+)?:bedrock-agentcore:[a-z0-9-]+:(aws|[0-9]{12}):code-interpreter(-custom)?/(aws\.codeinterpreter\.v1|[a-zA-Z][a-zA-Z0-9_]{0,47}-[a-zA-Z0-9]{10})`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

HarnessAgentCoreGatewayConfig

Configuration for AgentCore Gateway.

Contents

gatewayArn

The ARN of the desired AgentCore Gateway.

Type: String

Pattern: `arn:aws(|-cn|-us-gov):bedrock-agentcore:[a-z0-9-]{1,20}::[0-9]{12}:gateway/([0-9a-z][-]?)^{1,48}-[a-z0-9]{10}`

Required: Yes

outboundAuth

How harness authenticates to this Gateway. Defaults to AWS_IAM (SigV4) if omitted.

Type: [HarnessGatewayOutboundAuth](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

HarnessBedrockModelConfig

Configuration for an Amazon Bedrock model provider.

Contents

modelId

The Bedrock model ID.

Type: String

Required: Yes

additionalParams

Provider-specific parameters passed through to the model provider unchanged.

Type: JSON value

Required: No

apiFormat

The API format to use when calling the Bedrock provider.

Type: String

Valid Values: `converse_stream` | `responses` | `chat_completions`

Required: No

maxTokens

The maximum number of tokens to allow in the generated response per iteration.

Type: Integer

Valid Range: Minimum value of 1.

Required: No

temperature

The temperature to set when calling the model.

Type: Float

Valid Range: Minimum value of 0.0. Maximum value of 2.0.

Required: No

topP

The topP set when calling the model.

Type: Float

Valid Range: Minimum value of 0.0. Maximum value of 1.0.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

HarnessContentBlock

A content block within a message.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

reasoningContent

Model reasoning content.

Type: [HarnessReasoningContentBlock](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: No

text

Text content.

Type: String

Length Constraints: Minimum length of 1.

Required: No

toolResult

A tool execution result.

Type: [HarnessToolResultBlock](#) object

Required: No

toolUse

A tool use request from the model.

Type: [HarnessToolUseBlock](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

HarnessContentBlockDelta

A delta update to a content block.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

reasoningContent

A reasoning content delta.

Type: [HarnessReasoningContentBlockDelta](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: No

text

A text delta.

Type: String

Length Constraints: Minimum length of 1.

Required: No

toolResult

A tool result delta.

Type: Array of [HarnessToolResultBlockDelta](#) objects

Required: No

toolUse

A tool use input delta.

Type: [HarnessToolUseBlockDelta](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

HarnessContentBlockDeltaEvent

Event containing a delta update to a content block.

Contents

contentBlockIndex

The index of the content block being updated.

Type: Integer

Required: Yes

delta

The delta payload.

Type: [HarnessContentBlockDelta](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

HarnessContentBlockStart

The start payload for a content block.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

toolResult

Start of a tool result content block.

Type: [HarnessToolResultBlockStart](#) object

Required: No

toolUse

Start of a tool use content block.

Type: [HarnessToolUseBlockStart](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

HarnessContentBlockStartEvent

Event indicating the start of a content block.

Contents

contentBlockIndex

The index of the content block within the message.

Type: Integer

Required: Yes

start

The content block start payload.

Type: [HarnessContentBlockStart](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

HarnessContentBlockStopEvent

Event indicating the end of a content block.

Contents

contentBlockIndex

The index of the content block that ended.

Type: Integer

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

HarnessGatewayOutboundAuth

Authentication method for calling a Gateway.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

awslam

SigV4-sign requests using the agent's execution role.

Type: Structure

Required: No

none

No authentication.

Type: Structure

Required: No

oauth

OAuth 2.0 authentication via AgentCore Identity.

Type: [OAuthCredentialProvider](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)

- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

HarnessGeminiModelConfig

Configuration for a Google Gemini model provider. Requires an API key stored in AgentCore Identity.

Contents

apiKeyArn

The ARN of your Gemini API key on AgentCore Identity.

Type: String

Pattern: `arn:aws:bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:token-vault/[a-zA-Z0-9-\.]+/apikeycredentialprovider/[a-zA-Z0-9-\.]+`

Required: Yes

modelId

The Gemini model ID.

Type: String

Required: Yes

maxTokens

The maximum number of tokens to allow in the generated response per iteration.

Type: Integer

Valid Range: Minimum value of 1.

Required: No

temperature

The temperature to set when calling the model.

Type: Float

Valid Range: Minimum value of 0.0. Maximum value of 2.0.

Required: No

topK

The topK set when calling the model.

Type: Integer

Valid Range: Minimum value of 0. Maximum value of 500.

Required: No

topP

The topP set when calling the model.

Type: Float

Valid Range: Minimum value of 0.0. Maximum value of 1.0.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

HarnessInlineFunctionConfig

Configuration for an inline function tool. When the agent calls this tool, the tool call is returned to the caller for external execution.

Contents

description

Description of what the tool does, provided to the model.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 4096.

Required: Yes

inputSchema

JSON Schema describing the tool's input parameters.

Type: JSON value

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

HarnessLiteLlmModelConfig

Configuration for a LiteLLM model provider, enabling connection to third-party model providers.

Contents

modelId

The LiteLLM model identifier (e.g., "anthropic/claude-3-sonnet").

Type: String

Required: Yes

additionalParams

Provider-specific parameters passed through to the model provider unchanged.

Type: JSON value

Required: No

apiBase

The base URL for the model provider's API endpoint.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 16383.

Required: No

apiKeyArn

The ARN of the API key in AgentCore Identity for authenticating with the model provider.

Type: String

Pattern: `arn:aws:bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:token-vault/[a-zA-Z0-9-\.]+/apikeycredentialprovider/[a-zA-Z0-9-\.]+`

Required: No

maxTokens

The maximum number of tokens to allow in the generated response per iteration.

Type: Integer

Valid Range: Minimum value of 1.

Required: No

temperature

The temperature to set when calling the model.

Type: Float

Valid Range: Minimum value of 0.0. Maximum value of 2.0.

Required: No

topP

The topP set when calling the model.

Type: Float

Valid Range: Minimum value of 0.0. Maximum value of 1.0.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

HarnessMessage

A message in the conversation.

Contents

content

The content blocks of the message.

Type: Array of [HarnessContentBlock](#) objects

Required: Yes

role

The role of the message sender.

Type: String

Valid Values: user | assistant

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

HarnessMessageStartEvent

Event indicating the start of a message.

Contents

role

The role of the message sender.

Type: String

Valid Values: user | assistant

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

HarnessMessageStopEvent

Event indicating the end of a message.

Contents

stopReason

The reason the agent stopped generating.

Type: String

Valid Values: end_turn | tool_use | tool_result | max_tokens | stop_sequence
| content_filtered | malformed_model_output | malformed_tool_use
| interrupted | partial_turn | model_context_window_exceeded |
max_iterations_exceeded | max_output_tokens_exceeded | timeout_exceeded

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

HarnessMetadataEvent

Token usage and latency metrics for the invocation.

Contents

metrics

Latency metrics.

Type: [HarnessStreamMetrics](#) object

Required: Yes

usage

Token usage counts.

Type: [HarnessTokenUsage](#) object

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

HarnessModelConfiguration

Specification of which model to use.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

bedrockModelConfig

Configuration for an Amazon Bedrock model.

Type: [HarnessBedrockModelConfig](#) object

Required: No

geminiModelConfig

Configuration for a Google Gemini model.

Type: [HarnessGeminiModelConfig](#) object

Required: No

liteLlmModelConfig

The LiteLLM model configuration for connecting to third-party model providers.

Type: [HarnessLiteLlmModelConfig](#) object

Required: No

openAiModelConfig

Configuration for an OpenAI model.

Type: [HarnessOpenAiModelConfig](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

HarnessOpenAiModelConfig

Configuration for an OpenAI model provider. Requires an API key stored in AgentCore Identity.

Contents

apiKeyArn

The ARN of your OpenAI API key on AgentCore Identity.

Type: String

Pattern: `arn:aws:bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:token-vault/[a-zA-Z0-9-\.]+/apikeycredentialprovider/[a-zA-Z0-9-\.]+`

Required: Yes

modelId

The OpenAI model ID.

Type: String

Required: Yes

additionalParams

Provider-specific parameters passed through to the model provider unchanged.

Type: JSON value

Required: No

apiFormat

The API format to use when calling the OpenAI provider.

Type: String

Valid Values: `chat_completions` | `responses`

Required: No

maxTokens

The maximum number of tokens to allow in the generated response per iteration.

Type: Integer

Valid Range: Minimum value of 1.

Required: No

temperature

The temperature to set when calling the model.

Type: Float

Valid Range: Minimum value of 0.0. Maximum value of 2.0.

Required: No

topP

The topP set when calling the model.

Type: Float

Valid Range: Minimum value of 0.0. Maximum value of 1.0.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

HarnessReasoningContentBlock

Reasoning content from the model.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

reasoningText

The reasoning text.

Type: [HarnessReasoningTextBlock](#) object

Required: No

redactedContent

Redacted reasoning content.

Type: Base64-encoded binary data object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

HarnessReasoningContentBlockDelta

A delta update to a reasoning content block.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

redactedContent

Redacted reasoning content.

Type: Base64-encoded binary data object

Length Constraints: Minimum length of 0. Maximum length of 100000000.

Required: No

signature

Signature for the reasoning content.

Type: String

Required: No

text

Reasoning text delta.

Type: String

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

HarnessReasoningTextBlock

A block of reasoning text from the model.

Contents

text

The reasoning text.

Type: String

Required: Yes

signature

Signature for verifying the reasoning content.

Type: String

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

HarnessRemoteMcpConfig

Configuration for connecting to a remote MCP server.

Contents

url

URL of the MCP endpoint.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 16383.

Required: Yes

headers

Custom headers to include when connecting to the remote MCP server.

Type: String to string map

Key Length Constraints: Minimum length of 1. Maximum length of 16383.

Value Length Constraints: Minimum length of 1. Maximum length of 16383.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

HarnessSkill

A skill available to the agent.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

awsSkills

AWS Skills baked into the Harness's underlying Runtime.

Type: [HarnessSkillAwsSkillsSource](#) object

Required: No

git

A git repository containing the skill.

Type: [HarnessSkillGitSource](#) object

Required: No

path

The filesystem path to the skill definition.

Type: String

Length Constraints: Minimum length of 1.

Required: No

s3

An S3 source containing the skill.

Type: [HarnessSkillS3Source](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

HarnessSkillAwsSkillsSource

Passed to show that AWS Skills should be included.

Contents

paths

Optionally filter allowed skills with glob syntax, e.g., ['core-skills/*'].

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 4096.

Pattern: (`[^*?\\[\]]|*`)+

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

HarnessSkillGitAuth

Authentication configuration for accessing a private git repository.

Contents

credentialArn

The ARN of the credential in AgentCore Identity containing the password or personal access token.

Type: String

Pattern: `arn:aws:bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:token-vault/[a-zA-Z0-9-\.]+/apikeycredentialprovider/[a-zA-Z0-9-\.]+`

Required: Yes

username

Username for authentication. Defaults to 'oauth2' if not specified.

Type: String

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

HarnessSkillGitSource

A git repository source for a skill.

Contents

url

The HTTPS URL of the git repository.

Type: String

Length Constraints: Minimum length of 8.

Pattern: `https://[^\#@]+`

Required: Yes

auth

Authentication configuration for private repositories.

Type: [HarnessSkillGitAuth](#) object

Required: No

path

Subdirectory within the repository containing the skill.

Type: String

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

HarnessSkillS3Source

An S3 source for a skill.

Contents

uri

The S3 URI pointing to the skill directory (e.g., `s3://bucket/skills/my-skill/`).

Type: String

Length Constraints: Minimum length of 5.

Pattern: `s3://.*`

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

HarnessStreamMetrics

Latency metrics for the invocation.

Contents

latencyMs

The end-to-end latency of the invocation in milliseconds.

Type: Long

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

HarnessSystemContentBlock

A content block in the system prompt.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

text

The text content of the system prompt block.

Type: String

Length Constraints: Minimum length of 1.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

HarnessTokenUsage

Token usage counts for the invocation.

Contents

inputTokens

The number of input tokens consumed.

Type: Integer

Valid Range: Minimum value of 0.

Required: Yes

outputTokens

The number of output tokens generated.

Type: Integer

Valid Range: Minimum value of 0.

Required: Yes

totalTokens

The total number of tokens consumed.

Type: Integer

Valid Range: Minimum value of 0.

Required: Yes

cacheReadInputTokens

The number of input tokens read from cache.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

cacheWriteInputTokens

The number of input tokens written to cache.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

HarnessTool

A tool available to the agent loop.

Contents

type

The type of tool.

Type: String

Valid Values: `remote_mcp` | `agentcore_browser` | `agentcore_gateway` | `inline_function` | `agentcore_code_interpreter`

Required: Yes

config

Tool-specific configuration.

Type: [HarnessToolConfiguration](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: No

name

Unique name for the tool. If not provided, a name will be inferred or generated.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: `[a-zA-Z0-9_-]+`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

HarnessToolConfiguration

Configuration union for different tool types.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

agentCoreBrowser

Configuration for AgentCore Browser.

Type: [HarnessAgentCoreBrowserConfig](#) object

Required: No

agentCoreCodeInterpreter

Configuration for AgentCore Code Interpreter.

Type: [HarnessAgentCoreCodeInterpreterConfig](#) object

Required: No

agentCoreGateway

Configuration for AgentCore Gateway.

Type: [HarnessAgentCoreGatewayConfig](#) object

Required: No

inlineFunction

Configuration for an inline function tool.

Type: [HarnessInlineFunctionConfig](#) object

Required: No

remoteMcp

Configuration for remote MCP server.

Type: [HarnessRemoteMcpConfig](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

HarnessToolResultBlock

The result of a tool execution.

Contents

content

The content of the tool result.

Type: Array of [HarnessToolResultContentBlock](#) objects

Required: Yes

toolUseId

The tool use ID that this result corresponds to.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: `[a-zA-Z0-9_-]+`

Required: Yes

status

The status of the tool execution.

Type: String

Valid Values: `success` | `error`

Required: No

type

The type of tool use that produced this result.

Type: String

Valid Values: `tool_use` | `server_tool_use` | `mcp_tool_use`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

HarnessToolResultBlockDelta

A delta update to a tool result content block.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

json

A JSON tool result delta.

Type: JSON value

Required: No

text

A text tool result delta.

Type: String

Length Constraints: Minimum length of 1.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

HarnessToolResultBlockStart

Start payload for a tool result content block.

Contents

toolUseId

The tool use ID that this result corresponds to.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: `[a-zA-Z0-9_-]+`

Required: Yes

status

The status of the tool execution.

Type: String

Valid Values: `success` | `error`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

HarnessToolResultContentBlock

A content block within a tool result.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

json

JSON content.

Type: JSON value

Required: No

text

Text content.

Type: String

Length Constraints: Minimum length of 1.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

HarnessToolUseBlock

A tool use request from the model.

Contents

input

The JSON input to pass to the tool.

Type: JSON value

Required: Yes

name

The name of the tool to call.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: [a-zA-Z0-9_-]+

Required: Yes

toolUseId

The unique ID of this tool use.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: [a-zA-Z0-9_-]+

Required: Yes

serverName

The name of the MCP server providing this tool.

Type: String

Required: No

type

The type of tool use.

Type: String

Valid Values: `tool_use` | `server_tool_use` | `mcp_tool_use`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

HarnessToolUseBlockDelta

Delta payload for tool use input.

Contents

input

The partial JSON input for the tool call.

Type: String

Length Constraints: Minimum length of 1.

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

HarnessToolUseBlockStart

Start payload for a tool use content block.

Contents

name

The name of the tool being called.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: [a-zA-Z0-9_-]+

Required: Yes

toolUseId

The unique ID of this tool use.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: [a-zA-Z0-9_-]+

Required: Yes

serverName

The name of the MCP server providing this tool.

Type: String

Required: No

type

The type of tool use.

Type: String

Valid Values: tool_use | server_tool_use | mcp_tool_use

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

InlineGroundTruth

Inline ground truth data containing assertions, expected trajectories, and per-turn expected responses.

Contents

assertions

Assertions for evaluation, reuses common model EvaluationContentList.

Type: Array of [EvaluationContent](#) objects

Array Members: Minimum number of 1 item. Maximum number of 100 items.

Required: No

expectedTrajectory

The expected tool call sequence for trajectory evaluation.

Type: [EvaluationExpectedTrajectory](#) object

Required: No

turns

A list of per-turn ground truth data, each containing an input prompt and expected response.

Type: Array of [GroundTruthTurn](#) objects

Array Members: Minimum number of 1 item.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)

- [AWS SDK for Ruby V3](#)

InputContentBlock

A block of input content.

Contents

path

The path to the input content.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 100000000.

Required: Yes

blob

The binary input content.

Type: Base64-encoded binary data object

Length Constraints: Minimum length of 0. Maximum length of 100000000.

Required: No

text

The text input content.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 100000000.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)

- [AWS SDK for Ruby V3](#)

Insight

A reference to an insight analysis to run against sessions during batch evaluation. Insights provide deeper analysis beyond individual evaluator scores, including failure detection, user intent clustering, and execution summarization.

Contents

`insightId`

The unique identifier of the insight to run.

Type: String

Pattern: (Builtin\[a-zA-Z0-9._-]+|[a-zA-Z][a-zA-Z0-9-_{0,99}-[a-zA-Z0-9]{10})

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

InsightsFailureSignal

A signal indicating a detected failure within a span.

Contents

category

The failure category classification for this signal.

Type: String

Valid Values: execution-error-category-authentication | execution-error-category-resource-not-found | execution-error-category-service-errors | execution-error-category-rate-limiting | execution-error-category-formatting | execution-error-category-timeout | execution-error-category-resource-exhaustion | execution-error-category-environment | execution-error-category-tool-schema | task-instruction-category-non-compliance | task-instruction-category-problem-id | incorrect-actions-category-tool-selection | incorrect-actions-category-poor-information-retrieval | incorrect-actions-category-clarification | incorrect-actions-category-inappropriate-info-request | context-handling-error-category-context-handling-failures | hallucination-category-hall-capabilities | hallucination-category-hall-misunderstand | hallucination-category-hall-usage | hallucination-category-hall-history | hallucination-category-hall-params | hallucination-category-fabricate-tool-outputs | repetitive-behavior-category-repetition-tool | repetitive-behavior-category-repetition-info | repetitive-behavior-category-step-repetition | orchestration-related-errors-category-reasoning-mismatch | orchestration-related-errors-category-goal-deviation | orchestration-related-errors-category-premature-termination | orchestration-related-errors-category-unaware-termination | llm-output-category-nonsensical | configuration-mismatch-category-tool-definition | coding-use-case-specific-failure-types-category-edge-case-oversights | coding-use-case-specific-failure-types-category-dependency-issues | other

Required: Yes

confidence

The confidence score of the failure detection.

Type: Double

Required: Yes

evidence

The evidence supporting the failure detection.

Type: String

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

InvokeAgentRuntimeCommandRequestBody

The request body structure for the `InvokeAgentRuntimeCommand` operation, containing the command to execute and optional configuration parameters.

Contents

command

The shell command to execute on the agent runtime. This command is executed in the runtime environment and its output is streamed back to the caller.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 65536.

Required: Yes

timeout

The maximum duration in seconds to wait for the command to complete. If the command execution exceeds this timeout, it will be terminated. Default is 300 seconds. Minimum is 1 second. Maximum is 3600 seconds.

Type: Integer

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

InvokeAgentRuntimeCommandStreamOutput

The streaming output union for the `InvokeAgentRuntimeCommand` operation. This union delivers typed events: `contentStart` (first), `contentDelta` (middle), and `contentStop` (last).

Contents

`accessDeniedException`

Exception events for error streaming.

Type: Exception

HTTP Status Code: 403

Required: No

`chunk`

A response chunk containing command execution events such as content start, content delta, or content stop events.

Type: [ResponseChunk](#) object

Required: No

`internalServerErrorException`

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

Type: Exception

HTTP Status Code: 500

Required: No

`resourceNotFoundException`

The exception that occurs when the specified resource does not exist. This can happen when using an invalid identifier or when trying to access a resource that has been deleted.

Type: Exception

HTTP Status Code: 404

Required: No

runtimeClientError

The exception that occurs when there is an error in the runtime client. This can happen due to network issues, invalid configuration, or other client-side problems. Check the error message for specific details about the error.

Type: Exception

HTTP Status Code: 424

Required: No

serviceQuotaExceededException

The exception that occurs when the request would cause a service quota to be exceeded. Review your service quotas and either reduce your request rate or request a quota increase.

Type: Exception

HTTP Status Code: 402

Required: No

throttlingException

The exception that occurs when the request was denied due to request throttling. This happens when you exceed the allowed request rate for an operation. Reduce the frequency of requests or implement exponential backoff retry logic in your application.

Type: Exception

HTTP Status Code: 429

Required: No

validationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

Type: Exception

HTTP Status Code: 400

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

InvokeHarnessStreamOutput

The streaming events returned by a harness invocation.

Contents

contentBlockDelta

A delta update to the current content block.

Type: [HarnessContentBlockDeltaEvent](#) object

Required: No

contentBlockStart

Indicates the start of a new content block.

Type: [HarnessContentBlockStartEvent](#) object

Required: No

contentBlockStop

Indicates the end of the current content block.

Type: [HarnessContentBlockStopEvent](#) object

Required: No

internalServerErrorException

The exception that occurs when the service encounters an unexpected internal error. This is a temporary condition that will resolve itself with retries. We recommend implementing exponential backoff retry logic in your application.

Type: Exception

HTTP Status Code: 500

Required: No

messageStart

Indicates the start of a new message from the agent.

Type: [HarnessMessageStartEvent](#) object

Required: No

messageStop

Indicates the end of the current message.

Type: [HarnessMessageStopEvent](#) object

Required: No

metadata

Token usage and latency metrics for the invocation.

Type: [HarnessMetadataEvent](#) object

Required: No

runtimeClientError

An error returned by the runtime container during agent execution.

Type: Exception

HTTP Status Code: 424

Required: No

validationException

The exception that occurs when the input fails to satisfy the constraints specified by the service. Check the error message for details about which input parameter is invalid and correct your request.

Type: Exception

HTTP Status Code: 400

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

KeyPressArguments

Arguments for a key press action.

Contents

key

The key name to press (for example, `enter`, `tab`, `escape`).

Type: String

Required: Yes

presses

The number of times to press the key. Valid range: 1–100. Defaults to 1.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

KeyPressResult

The result of a key press action.

Contents

status

The status of the action execution.

Type: String

Valid Values: SUCCESS | FAILED

Required: Yes

error

The error message. Present only when the action failed.

Type: String

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

KeyShortcutArguments

Arguments for a key shortcut action.

Contents

keys

The key combination to press (for example, ["ctrl", "s"]). Maximum 5 keys.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 5 items.

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

KeyShortcutResult

The result of a key shortcut action.

Contents

status

The status of the action execution.

Type: String

Valid Values: SUCCESS | FAILED

Required: Yes

error

The error message. Present only when the action failed.

Type: String

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

KeyTypeArguments

Arguments for a key type action.

Contents

text

The text string to type. Maximum length: 10,000 characters.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 10000.

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

KeyTypeResult

The result of a key type action.

Contents

status

The status of the action execution.

Type: String

Valid Values: SUCCESS | FAILED

Required: Yes

error

The error message. Present only when the action failed.

Type: String

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

LeftExpression

Left expression of the event metadata filter.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

metadataKey

Key associated with the metadata in an event.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 128.

Pattern: `[a-zA-Z0-9\s._:/=+@-]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

LinkedAccount

Represents different linked accounts that can be linked to an embedded wallet. Supports email, SMS, JWT, and OAuth2 approaches.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

developerJwt

Developer JWT linked account with key ID and subject.

Type: [LinkedAccountDeveloperJwt](#) object

Required: No

email

Email-based linked account.

Type: [LinkedAccountEmail](#) object

Required: No

oAuth2

OAuth2 provider linked account (Google, Apple, X, Telegram, GitHub).

Type: [LinkedAccountOAuth2](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: No

sms

SMS-based linked account using phone number.

Type: [LinkedAccountSms](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

LinkedAccountDeveloperJwt

Authentication method using JWT with key ID and subject claims.

Contents

kid

The key ID (kid) from the JWT header. Identifies which key was used to sign the JWT.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[a-zA-Z0-9_-]{1,255}`

Required: Yes

sub

The subject (sub) claim from the JWT payload. Identifies the principal that is the subject of the JWT.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

LinkedAccountEmail

Linked account using an email address.

Contents

emailAddress

The email address used for the linked account. Must be a valid email format.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 254.

Pattern: `[a-zA-Z0-9._%+-]+@[a-zA-Z0-9.-]+\.[a-zA-Z]{2,}`

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

LinkedAccountOAuth2

Authentication method using OAuth2 providers. Supports Google, Apple, X, Telegram, and GitHub providers.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

apple

Apple OAuth2 authentication.

Type: [OAuth2Authentication](#) object

Required: No

github

GitHub OAuth2 authentication.

Type: [OAuth2Authentication](#) object

Required: No

google

Google OAuth2 authentication.

Type: [OAuth2Authentication](#) object

Required: No

telegram

Telegram OAuth2 authentication.

Type: [OAuth2Authentication](#) object

Required: No

x

X (formerly Twitter) OAuth2 authentication.

Type: [OAuth2Authentication](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

LinkedAccountSms

Linked account using a phone number in E.164 format.

Contents

phoneNumber

The phone number in E.164 format (e.g., +1234567890).

Type: String

Length Constraints: Minimum length of 3. Maximum length of 16.

Pattern: `\+[1-9]\d{1,14}`

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

LiveViewStream

The configuration for a stream that provides a visual representation of a browser session in Amazon Bedrock AgentCore. This stream enables agents to observe the current state of the browser, including rendered web pages, visual elements, and the results of interactions.

Contents

`streamEndpoint`

The endpoint URL for the live view stream. This URL is used to establish a connection to receive visual updates from the browser session.

Type: String

Length Constraints: Minimum length of 10. Maximum length of 512.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

McpDescriptor

The MCP (Model Context Protocol) descriptor configuration for a registry record. Contains the server definition and tools definition.

Contents

server

The MCP server definition that describes the server configuration.

Type: [ServerDefinition](#) object

Required: Yes

tools

The MCP tools definition that describes the available tools.

Type: [ToolsDefinition](#) object

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

MemoryContent

Contains the content of a memory record.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

text

The text content of the memory record.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 16000.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

MemoryMetadataFilterExpression

Filters to apply to metadata associated with a memory. Specify the metadata key and value in the `left` and `right` fields and use the `operator` field to define the relationship to match.

Contents

left

The metadata key to evaluate.

Type: [MemoryRecordLeftExpression](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: Yes

operator

The relationship between the metadata key and value to match when applying the metadata filter.

Type: String

Valid Values: EQUALS_TO | EXISTS | NOT_EXISTS | BEFORE | AFTER | CONTAINS | GREATER_THAN | GREATER_THAN_OR_EQUALS | LESS_THAN | LESS_THAN_OR_EQUALS

Required: Yes

right

The value to compare against. Required for all operators except EXISTS and NOT_EXISTS.

Type: [MemoryRecordRightExpression](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

MemoryRecord

Contains information about a memory record in an AgentCore Memory resource.

Contents

content

The content of the memory record.

Type: [MemoryContent](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: Yes

createdAt

The timestamp when the memory record was created.

Type: Timestamp

Required: Yes

memoryRecordId

The unique identifier of the memory record.

Type: String

Length Constraints: Minimum length of 40. Maximum length of 50.

Pattern: mem-[a-zA-Z0-9-]*

Required: Yes

memoryStrategyId

The identifier of the memory strategy associated with this record.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 100.

Pattern: [a-zA-Z0-9][a-zA-Z0-9-]*

Required: Yes

namespaces

The namespaces associated with this memory record. Namespaces help organize and categorize memory records.

Type: Array of strings

Array Members: Minimum number of 0 items. Maximum number of 1 item.

Length Constraints: Minimum length of 1. Maximum length of 1024.

Pattern: `[a-zA-Z0-9/*][a-zA-Z0-9_/*]*(?::[a-zA-Z0-9_/*]+)*[a-zA-Z0-9_/*]*`

Required: Yes

metadata

A map of metadata key-value pairs associated with a memory record.

Type: String to [MemoryRecordMetadataValue](#) object map

Map Entries: Maximum number of 20 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Key Pattern: `[a-zA-Z0-9\s._:/=+@-]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

MemoryRecordCreateInput

Input structure to create a new memory record.

Contents

content

The content to be stored within the memory record.

Type: [MemoryContent](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: Yes

namespaces

A list of namespace identifiers that categorize or group the memory record.

Type: Array of strings

Array Members: Minimum number of 0 items. Maximum number of 1 item.

Length Constraints: Minimum length of 1. Maximum length of 1024.

Pattern: `[a-zA-Z0-9/*][a-zA-Z0-9-_/]*(:|[a-zA-Z0-9-_/]*+)*[a-zA-Z0-9-_/]*`

Required: Yes

requestIdentifier

A client-provided identifier for tracking this specific record creation request.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 80.

Pattern: `[a-zA-Z0-9_-]+`

Required: Yes

timestamp

Time at which the memory record was created.

Type: Timestamp

Required: Yes

memoryStrategyId

The ID of the memory strategy that defines how this memory record is grouped.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 100.

Pattern: [a-zA-Z0-9][a-zA-Z0-9-_*]

Required: No

metadata

Metadata key-value pairs to be stored with the memory record.

Type: String to [MemoryRecordMetadataValue](#) object map

Map Entries: Maximum number of 20 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Key Pattern: [a-zA-Z0-9\s._:/=+@-]*

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

MemoryRecordDeleteInput

Input structure to delete an existing memory record.

Contents

memoryRecordId

The unique ID of the memory record to be deleted.

Type: String

Length Constraints: Minimum length of 40. Maximum length of 50.

Pattern: mem-[a-zA-Z0-9-_*]

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

MemoryRecordLeftExpression

The left-hand side of a memory record metadata filter expression.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

metadataKey

The metadata key to filter on.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 128.

Pattern: `[a-zA-Z0-9\s._:/=+@-]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

MemoryRecordMetadataValue

The value of a memory record metadata entry.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

dateTimeValue

A timestamp value in ISO 8601 UTC format.

Type: Timestamp

Required: No

numberValue

A numeric value.

Type: Double

Required: No

stringListValue

A list of string values.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 5 items.

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: `[a-zA-Z0-9\s._:/=+@-]*`

Required: No

stringValue

A string value.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: [a-zA-Z0-9\s._:/=+@-]*

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

MemoryRecordOutput

Output information returned after processing a memory record operation.

Contents

memoryRecordId

The unique ID associated to the memory record.

Type: String

Length Constraints: Minimum length of 40. Maximum length of 50.

Pattern: mem-[a-zA-Z0-9-]*

Required: Yes

status

The status of the memory record operation (e.g., SUCCEEDED, FAILED).

Type: String

Valid Values: SUCCEEDED | FAILED

Required: Yes

errorCode

The error code returned when the memory record operation fails.

Type: Integer

Required: No

errorMessage

A human-readable error message describing why the memory record operation failed.

Type: String

Required: No

requestIdentifier

The client-provided identifier that was used to track this record operation.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 80.

Pattern: [a-zA-Z0-9_-]+

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

MemoryRecordRightExpression

The right-hand side of a memory record metadata filter expression.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

metadataValue

The metadata value to compare against.

Type: [MemoryRecordMetadataValue](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

MemoryRecordSummary

Contains summary information about a memory record.

Contents

content

The content of the memory record.

Type: [MemoryContent](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: Yes

createdAt

The timestamp when the memory record was created.

Type: Timestamp

Required: Yes

memoryRecordId

The unique identifier of the memory record.

Type: String

Length Constraints: Minimum length of 40. Maximum length of 50.

Pattern: mem-[a-zA-Z0-9-]*

Required: Yes

memoryStrategyId

The identifier of the memory strategy associated with this record.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 100.

Pattern: [a-zA-Z0-9][a-zA-Z0-9-]*

Required: Yes

namespaces

The namespaces associated with this memory record.

Type: Array of strings

Array Members: Minimum number of 0 items. Maximum number of 1 item.

Length Constraints: Minimum length of 1. Maximum length of 1024.

Pattern: `[a-zA-Z0-9/*][a-zA-Z0-9-_/]*(?::[a-zA-Z0-9-_/]*+)*[a-zA-Z0-9-_/]*`

Required: Yes

metadata

A map of metadata key-value pairs associated with a memory record.

Type: String to [MemoryRecordMetadataValue](#) object map

Map Entries: Maximum number of 20 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Key Pattern: `[a-zA-Z0-9\s._:/=+@-]*`

Required: No

score

The relevance score of the memory record when returned as part of a search result. Higher values indicate greater relevance to the search query.

Type: Double

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

MemoryRecordUpdateInput

Input structure to update an existing memory record.

Contents

memoryRecordId

The unique ID of the memory record to be updated.

Type: String

Length Constraints: Minimum length of 40. Maximum length of 50.

Pattern: mem-[a-zA-Z0-9-]*

Required: Yes

timestamp

Time at which the memory record was updated

Type: Timestamp

Required: Yes

content

The content to be stored within the memory record.

Type: [MemoryContent](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: No

memoryStrategyId

The updated ID of the memory strategy that defines how this memory record is grouped.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 100.

Pattern: [a-zA-Z0-9][a-zA-Z0-9-]*

Required: No

metadata

Metadata key-value pairs to be stored with the memory record.

Type: String to [MemoryRecordMetadataValue](#) object map

Map Entries: Maximum number of 20 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Key Pattern: `[a-zA-Z0-9\s._:/=+@-]*`

Required: No

namespaces

The updated list of namespace identifiers for categorizing the memory record.

Type: Array of strings

Array Members: Minimum number of 0 items. Maximum number of 1 item.

Length Constraints: Minimum length of 1. Maximum length of 1024.

Pattern: `[a-zA-Z0-9/*][a-zA-Z0-9-_/]*(?::[a-zA-Z0-9-_/]*+)*[a-zA-Z0-9-_/]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

MessageMetadata

Metadata information associated with this message.

Contents

eventId

The identifier of the event associated with this message.

Type: String

Required: Yes

messageIndex

The position of this message within that event's ordered list of messages.

Type: Integer

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

MetadataValue

Value associated with the eventMetadata key.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

stringValue

Value associated with the eventMetadata key.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 256.

Pattern: `[a-zA-Z0-9\s._:/=+@-]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

MouseClickArguments

Arguments for a mouse click action.

Contents

x

The X coordinate on screen where the click occurs.

Type: Integer

Required: Yes

y

The Y coordinate on screen where the click occurs.

Type: Integer

Required: Yes

button

The mouse button to use. Defaults to LEFT.

Type: String

Valid Values: LEFT | RIGHT | MIDDLE

Required: No

clickCount

The number of clicks to perform. Valid range: 1–10. Defaults to 1.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 10.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

MouseClickedResult

The result of a mouse click action.

Contents

status

The status of the action execution.

Type: String

Valid Values: SUCCESS | FAILED

Required: Yes

error

The error message. Present only when the action failed.

Type: String

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

MouseDownArguments

Arguments for a mouse drag action.

Contents

endX

The ending X coordinate for the drag.

Type: Integer

Required: Yes

endY

The ending Y coordinate for the drag.

Type: Integer

Required: Yes

startX

The starting X coordinate for the drag.

Type: Integer

Required: Yes

startY

The starting Y coordinate for the drag.

Type: Integer

Required: Yes

button

The mouse button to use for the drag. Defaults to LEFT.

Type: String

Valid Values: LEFT | RIGHT | MIDDLE

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

MouseDownResult

The result of a mouse drag action.

Contents

status

The status of the action execution.

Type: String

Valid Values: SUCCESS | FAILED

Required: Yes

error

The error message. Present only when the action failed.

Type: String

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

MouseMoveArguments

Arguments for a mouse move action.

Contents

x

The target X coordinate on screen.

Type: Integer

Required: Yes

y

The target Y coordinate on screen.

Type: Integer

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

MouseMoveResult

The result of a mouse move action.

Contents

status

The status of the action execution.

Type: String

Valid Values: SUCCESS | FAILED

Required: Yes

error

The error message. Present only when the action failed.

Type: String

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

MouseScrollArguments

Arguments for a mouse scroll action.

Contents

x

The X coordinate on screen where the scroll occurs.

Type: Integer

Required: Yes

y

The Y coordinate on screen where the scroll occurs.

Type: Integer

Required: Yes

deltaX

The horizontal scroll delta. Valid range: -1000 to 1000.

Type: Integer

Valid Range: Minimum value of -1000. Maximum value of 1000.

Required: No

deltaY

The vertical scroll delta. Valid range: -1000 to 1000. Negative values scroll down.

Type: Integer

Valid Range: Minimum value of -1000. Maximum value of 1000.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

MouseScrollResult

The result of a mouse scroll action.

Contents

status

The status of the action execution.

Type: String

Valid Values: SUCCESS | FAILED

Required: Yes

error

The error message. Present only when the action failed.

Type: String

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

OAuth2Authentication

OAuth2 authentication information for third-party providers.

Contents

sub

The subject (sub) claim from the OAuth2 provider. Uniquely identifies the user at the provider.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Required: Yes

emailAddress

The email address from the OAuth2 provider.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 254.

Pattern: `[a-zA-Z0-9._%+-]+@[a-zA-Z0-9.-]+\.[a-zA-Z]{2,}`

Required: No

name

The user's name from the OAuth2 provider.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Required: No

username

The username from the OAuth2 provider.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

OAuthCredentialProvider

Configuration for an OAuth 2.0 credential provider used to authenticate tool calls.

Contents

providerArn

The ARN of the OAuth 2.0 credential provider in AgentCore Identity.

Type: String

Pattern: `arn:(^[^:]*):([^\:]*):([^\:]*):([0-9]{12})?:(.+)`

Required: Yes

scopes

The OAuth 2.0 scopes to request when obtaining an access token.

Type: Array of strings

Array Members: Minimum number of 0 items. Maximum number of 100 items.

Length Constraints: Minimum length of 1. Maximum length of 64.

Required: Yes

customParameters

Additional custom parameters to include in the OAuth 2.0 token request.

Type: String to string map

Map Entries: Maximum number of 10 items.

Key Length Constraints: Minimum length of 1. Maximum length of 256.

Value Length Constraints: Minimum length of 1. Maximum length of 2048.

Required: No

defaultReturnUrl

The default return URL for the OAuth 2.0 authorization flow.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: `\w+ : (\V?\V?)[^\s]+`

Required: No

grantType

The OAuth 2.0 grant type to use for authentication.

Type: String

Valid Values: CLIENT_CREDENTIALS | AUTHORIZATION_CODE | TOKEN_EXCHANGE

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

OnlineEvaluationConfigSource

A reference to an existing online evaluation configuration to use as the data source for batch evaluation.

Contents

onlineEvaluationConfigArn

The Amazon Resource Name (ARN) of the online evaluation configuration to use as the session source.

Type: String

Pattern: `arn:aws[a-zA-Z-]*:bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:online-evaluation-config\[a-zA-Z][a-zA-Z0-9-_\]{0,99}-[a-zA-Z0-9]{10}`

Required: Yes

timeRange

Optional session filter configuration to narrow down which sessions from the online evaluation configuration to include.

Type: [SessionFilterConfig](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

OutputConfig

Output destination configuration.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

cloudWatchConfig

The CloudWatch Logs configuration for writing evaluation results.

Type: [CloudWatchOutputConfig](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

PayloadType

Contains the payload content for an event.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

blob

The binary content of the payload.

Type: JSON value

Required: No

conversational

The conversational content of the payload.

Type: [Conversational](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

PaymentInput

The payment input details, which vary by payment type.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

cryptoX402

Input for a crypto X402 payment.

Type: [CryptoX402PaymentInput](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

PaymentInstrument

Represents a payment instrument.

Contents

createdAt

The timestamp when this payment instrument was created.

Type: Timestamp

Required: Yes

paymentConnectorId

The ID of the payment connector associated with this instrument.

Type: String

Length Constraints: Minimum length of 12. Maximum length of 211.

Pattern: (`[0-9a-z][-]?{1,100}`)-`[0-9a-z]{10}`

Required: Yes

paymentInstrumentDetails

The details specific to the payment instrument type.

Type: [PaymentInstrumentDetails](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: Yes

paymentInstrumentId

The unique identifier for this payment instrument.

Type: String

Length Constraints: Fixed length of 34.

Pattern: `payment-instrument-[0-9a-zA-Z-]{15}`

Required: Yes

paymentInstrumentType

The type of payment instrument (e.g., EMBEDDED_CRYPTOWALLET).

Type: String

Valid Values: EMBEDDED_CRYPTOWALLET

Required: Yes

paymentManagerArn

The ARN of the payment manager that owns this payment instrument.

Type: String

Length Constraints: Minimum length of 66. Maximum length of 2048.

Pattern: `arn:(aws|aws-[a-z0-9-]+):bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:payment-manager/[a-z0-9]([a-z0-9-]{0,47}[a-z0-9])?-[a-z0-9]{10}`

Required: Yes

status

The current status of this payment instrument.

Type: String

Valid Values: INITIATED | ACTIVE | FAILED | DELETED

Required: Yes

updatedAt

The timestamp when this payment instrument was last updated.

Type: Timestamp

Required: Yes

userId

The user ID associated with this payment instrument.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 120.

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

PaymentInstrumentDetails

Details specific to the instrument type.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

embeddedCryptoWallet

Embedded crypto wallet managed directly by end user.

Type: [EmbeddedCryptoWallet](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

PaymentInstrumentSummary

Summary of a payment instrument for list operations.

Contents

createdAt

The timestamp when this payment instrument was created.

Type: Timestamp

Required: Yes

paymentConnectorId

The ID of the payment connector associated with this instrument.

Type: String

Length Constraints: Minimum length of 12. Maximum length of 211.

Pattern: (`[0-9a-z][-]?{1,100}`)-`[0-9a-z]{10}`

Required: Yes

paymentInstrumentId

The unique identifier for this payment instrument.

Type: String

Length Constraints: Fixed length of 34.

Pattern: `payment-instrument-[0-9a-zA-Z-]{15}`

Required: Yes

paymentInstrumentType

The type of payment instrument (e.g., `EMBEDDED_CRYPTOWALLET`).

Type: String

Valid Values: EMBEDDED_CRYPTOWALLET

Required: Yes

paymentManagerArn

The ARN of the payment manager that owns this payment instrument.

Type: String

Length Constraints: Minimum length of 66. Maximum length of 2048.

Pattern: `arn:(aws|aws-[a-z0-9-]+):bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:payment-manager/[a-z0-9]([a-z0-9]{0,47}[a-z0-9])?-[a-z0-9]{10}`

Required: Yes

status

The current status of this payment instrument.

Type: String

Valid Values: INITIATED | ACTIVE | FAILED | DELETED

Required: Yes

updatedAt

The timestamp when this payment instrument was last updated.

Type: Timestamp

Required: Yes

userId

The user ID associated with this payment instrument.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 120.

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

PaymentOutput

The payment output details, which vary by payment type.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

cryptoX402

Output from a crypto X402 payment.

Type: [CryptoX402PaymentOutput](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

PaymentSession

A payment session for managing payment transactions.

Contents

createdAt

The timestamp when the session was created.

Type: Timestamp

Required: Yes

expiryTimeInMinutes

The session expiry time in minutes.

Type: Integer

Required: Yes

paymentManagerArn

The ARN of the payment manager that owns this session.

Type: String

Length Constraints: Minimum length of 66. Maximum length of 2048.

Pattern: `arn:(aws|aws-[a-z0-9-]+):bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:payment-manager/[a-z0-9]([a-z0-9-]{0,47}[a-z0-9])?-[a-z0-9]{10}`

Required: Yes

paymentSessionId

The unique identifier of the payment session.

Type: String

Length Constraints: Fixed length of 31.

Pattern: `payment-session-[0-9a-zA-Z-]{15}`

Required: Yes

updatedAt

The timestamp when the session was last updated.

Type: Timestamp

Required: Yes

userId

The user ID associated with this session.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 120.

Required: Yes

availableLimits

The current available spending limits.

Type: [AvailableLimits](#) object

Required: No

limits

The spending limits for the payment session.

Type: [SessionLimits](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

PaymentSessionSummary

Summary information about a payment session.

Contents

createdAt

The timestamp when the session was created.

Type: Timestamp

Required: Yes

expiryTimeInMinutes

The session expiry time in minutes.

Type: Integer

Required: Yes

paymentManagerArn

The ARN of the payment manager that owns this session.

Type: String

Length Constraints: Minimum length of 66. Maximum length of 2048.

Pattern: `arn:(aws|aws-[a-z0-9-]+):bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:payment-manager/[a-z0-9]([a-z0-9-]{0,47}[a-z0-9])?-[a-z0-9]{10}`

Required: Yes

paymentSessionId

The unique identifier of the payment session.

Type: String

Length Constraints: Fixed length of 31.

Pattern: `payment-session-[0-9a-zA-Z-]{15}`

Required: Yes

updatedAt

The timestamp when the session was last updated.

Type: Timestamp

Required: Yes

userId

The user ID associated with this session.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 120.

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

PaymentTokenRequestInput

Vendor-specific token request configuration.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

coinbaseCdpTokenRequest

The Coinbase CDP token request.

Type: [CoinbaseCdpTokenRequestInput](#) object

Required: No

stripePrivyTokenRequest

The Stripe Privy token request.

Type: [StripePrivyTokenRequestInput](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

PaymentTokenResponseOutput

Vendor-specific token response configuration.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

coinbaseCdpTokenResponse

The Coinbase CDP token response.

Type: [CoinbaseCdpTokenResponseOutput](#) object

Required: No

stripePrivyTokenResponse

The Stripe Privy token response.

Type: [StripePrivyTokenResponseOutput](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

PerVariantOnlineEvaluationConfig

An online evaluation configuration associated with a specific A/B test variant.

Contents

name

The name of the variant this evaluation configuration applies to.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2.

Pattern: (C|T1)

Required: Yes

onlineEvaluationConfigArn

The Amazon Resource Name (ARN) of the online evaluation configuration for this variant.

Type: String

Pattern: arn:aws[a-zA-Z-]*:bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:online-evaluation-config\[a-zA-Z][a-zA-Z0-9-_\]{0,99}-[a-zA-Z0-9]{10}

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Proxy

Union type representing different proxy configurations. Currently supports external customer-managed proxies.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

externalProxy

Configuration for an external customer-managed proxy server.

Type: [ExternalProxy](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ProxyBypass

Configuration for domains that should bypass all proxies and connect directly to the internet. These bypass rules take precedence over all proxy routing rules.

Contents

domainPatterns

Array of domain patterns that should bypass the proxy. Supports `.amazonaws.com` for subdomain matching or `amazonaws.com` for exact domain matching. Requests to these domains connect directly without using any proxy. Maximum 253 characters per pattern.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 100 items.

Length Constraints: Minimum length of 1. Maximum length of 253.

Pattern: `(\.)?[a-zA-Z0-9]([a-zA-Z0-9-]{0,61}[a-zA-Z0-9])?(\.[a-zA-Z0-9]([a-zA-Z0-9-]{0,61}[a-zA-Z0-9])?)*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ProxyConfiguration

Configuration for routing browser traffic through customer-managed proxy servers. Supports 1-5 proxy servers for domain-based routing and proxy bypass rules.

Contents

proxies

An array of 1-5 proxy server configurations for domain-based routing. Each proxy can specify which domains it handles via `domainPatterns`, enabling flexible routing of different traffic through different proxies based on destination domain.

Type: Array of [Proxy](#) objects

Array Members: Minimum number of 1 item. Maximum number of 5 items.

Required: Yes

bypass

Optional configuration for domains that should bypass all proxies and connect directly to their destination, like the internet. Takes precedence over all proxy routing rules.

Type: [ProxyBypass](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ProxyCredentials

Union type representing different proxy authentication methods. Currently supports HTTP Basic Authentication (username and password).

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

basicAuth

HTTP Basic Authentication credentials (username and password) stored in AWS Secrets Manager.

Type: [BasicAuth](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

RecommendationConfig

The configuration for a recommendation, varying by recommendation type.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

systemPromptRecommendationConfig

The configuration for a system prompt recommendation.

Type: [SystemPromptRecommendationConfig](#) object

Required: No

toolDescriptionRecommendationConfig

The configuration for a tool description recommendation.

Type: [ToolDescriptionRecommendationConfig](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

RecommendationEvaluationConfig

The evaluation configuration for assessing recommendation quality.

Contents

evaluators

The list of evaluators to use for assessing recommendation quality.

Type: Array of [RecommendationEvaluatorReference](#) objects

Array Members: Fixed number of 1 item.

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

RecommendationEvaluatorReference

A reference to an evaluator used for recommendation assessment.

Contents

evaluatorArn

The Amazon Resource Name (ARN) of the evaluator.

Type: String

Pattern: `arn:aws[a-zA-Z-]*:bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:evaluator`
`\/[a-zA-Z][a-zA-Z0-9-]{0,99}-[a-zA-Z0-9]{10}$|^arn:aws[a-zA-Z-]*:bedrock-agentcore::evaluator/Builtin.[a-zA-Z0-9_-]+`

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

RecommendationResult

The result of a recommendation, containing the optimized output.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

systemPromptRecommendationResult

The result of a system prompt recommendation.

Type: [SystemPromptRecommendationResult](#) object

Required: No

toolDescriptionRecommendationResult

The result of a tool description recommendation.

Type: [ToolDescriptionRecommendationResult](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

RecommendationResultConfigurationBundle

A configuration bundle reference in a recommendation result.

Contents

bundleArn

The Amazon Resource Name (ARN) of the configuration bundle.

Type: String

Pattern: `arn:aws[a-zA-Z-]*:bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:configuration-bundle/[a-zA-Z][a-zA-Z0-9-_{0,99}]-[a-zA-Z0-9]{10}`

Required: Yes

versionId

The version identifier of the configuration bundle containing the recommendation.

Type: String

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

RecommendationSummary

Summary information about a recommendation.

Contents

createdAt

The timestamp when the recommendation was created.

Type: Timestamp

Required: Yes

name

The name of the recommendation.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 100.

Pattern: `[a-zA-Z][a-zA-Z0-9_-]{0,47}`

Required: Yes

recommendationArn

The Amazon Resource Name (ARN) of the recommendation.

Type: String

Pattern: `arn:aws[a-zA-Z-]*:bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:recommendation/[0-9a-zA-Z_-]{1,48}-[0-9A-Z]{10}`

Required: Yes

recommendationId

The unique identifier of the recommendation.

Type: String

Pattern: `[0-9a-zA-Z_-]{1,48}-[0-9A-Z]{10}`

Required: Yes

status

The current status of the recommendation.

Type: String

Valid Values: PENDING | IN_PROGRESS | COMPLETED | FAILED | DELETING

Required: Yes

type

The type of recommendation.

Type: String

Valid Values: SYSTEM_PROMPT_RECOMMENDATION | TOOL_DESCRIPTION_RECOMMENDATION

Required: Yes

updatedAt

The timestamp when the recommendation was last updated.

Type: Timestamp

Required: Yes

description

The description of the recommendation.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 4096.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

RegistryRecordSummary

Summary information about a registry record.

Contents

createdAt

The date and time when the registry record was created.

Type: Timestamp

Required: Yes

descriptors

The descriptor configurations for this registry record.

Type: [Descriptors](#) object

Required: Yes

descriptorType

The type of descriptor associated with this registry record.

Type: String

Valid Values: MCP | A2A | CUSTOM | AGENT_SKILLS

Required: Yes

name

The name of the registry record.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[a-zA-Z0-9][a-zA-Z0-9_-\.\|]*`

Required: Yes

recordArn

The Amazon Resource Name (ARN) of the registry record.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: `arn:aws(-[^\:]+)?:bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:registry/[a-zA-Z0-9]{12,16}/record/[a-zA-Z0-9]{12}`

Required: Yes

recordId

The unique identifier of the registry record.

Type: String

Length Constraints: Fixed length of 12.

Pattern: `[a-zA-Z0-9]{12}`

Required: Yes

registryArn

The Amazon Resource Name (ARN) of the registry that this record belongs to.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: `arn:aws(-[^\:]+)?:bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:registry/[a-zA-Z0-9]{12,16}`

Required: Yes

status

The current status of the registry record.

Type: String

Valid Values: DRAFT | PENDING_APPROVAL | APPROVED | REJECTED | DEPRECATED

Required: Yes

updatedAt

The date and time when the registry record was last updated.

Type: Timestamp

Required: Yes

version

The version of the registry record.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: [a-zA-Z0-9.-]+

Required: Yes

description

A description of the registry record.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 4096.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ResourceContent

Contains information about resource content.

Contents

type

The type of resource content.

Type: String

Valid Values: `text` | `blob`

Required: Yes

blob

The binary resource content.

Type: Base64-encoded binary data object

Required: No

contentType

The MIME type of the resource content.

Type: String

Required: No

text

The text resource content.

Type: String

Required: No

uri

The URI of the resource content.

Type: String

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ResourceLocation

The location of the browser extension.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

s3

The Amazon S3 location of the resource. Use this when the resource is stored in an Amazon S3 bucket.

Type: [S3Location](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ResponseChunk

A structure representing a response chunk that contains exactly one of the possible event types: `contentStart`, `contentDelta`, or `contentStop`.

Contents

`contentDelta`

An event containing incremental output (stdout or stderr) from the command execution. These are the middle chunks.

Type: [ContentDeltaEvent](#) object

Required: No

`contentStart`

An event indicating the start of content streaming from the command execution. This is the first chunk received.

Type: [ContentStartEvent](#) object

Required: No

`contentStop`

An event indicating the completion of the command execution, including the exit code and final status. This is the last chunk received.

Type: [ContentStopEvent](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)

- [AWS SDK for Ruby V3](#)

RightExpression

Right expression of the eventMetadatafilter.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

metadataValue

Value associated with the key in eventMetadata.

Type: [MetadataValue](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

RootCauseCluster

A cluster of similar root causes identified within a failure subcategory.

Contents

affectedSessionCount

The number of sessions affected by this root cause.

Type: Integer

Required: Yes

affectedSessions

The list of sessions affected by this root cause.

Type: Array of [AffectedSession](#) objects

Array Members: Minimum number of 0 items.

Required: Yes

clusterId

The unique identifier of the root cause cluster.

Type: Integer

Required: Yes

name

The name of the root cause cluster.

Type: String

Required: Yes

recommendation

The recommended fix for this root cause.

Type: String

Required: Yes

rootCause

The root cause explanation for this cluster of failures.

Type: String

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

S3Location

The Amazon S3 location configuration of a resource.

Contents

bucket

The name of the Amazon S3 bucket where the resource is stored.

Type: String

Length Constraints: Minimum length of 3. Maximum length of 63.

Pattern: `[a-z0-9][a-z0-9.-]*[a-z0-9]`

Required: Yes

prefix

The name of the Amazon S3 prefix/key where the resource is stored.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1024.

Required: Yes

versionId

The name of the Amazon S3 version ID where the resource is stored (Optional).

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1024.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ScreenshotArguments

Arguments for a screenshot action.

Contents

format

The image format for the screenshot. Defaults to PNG.

Type: String

Valid Values: PNG

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ScreenshotResult

The result of a screenshot action.

Contents

status

The status of the action execution.

Type: String

Valid Values: SUCCESS | FAILED

Required: Yes

data

The base64-encoded image data. Present only when the action succeeded.

Type: Base64-encoded binary data object

Required: No

error

The error message. Present only when the action failed.

Type: String

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

SearchCriteria

Contains search criteria for retrieving memory records.

Contents

searchQuery

The search query to use for finding relevant memory records.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 10000.

Required: Yes

memoryStrategyId

The memory strategy identifier to filter memory records by.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 100.

Pattern: `[a-zA-Z0-9][a-zA-Z0-9-_*]`

Required: No

metadataFilters

Filters to apply to metadata associated with a memory.

Type: Array of [MemoryMetadataFilterExpression](#) objects

Array Members: Minimum number of 1 item. Maximum number of 5 items.

Required: No

topK

The maximum number of top-scoring memory records to return. This value is used for semantic search ranking.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

SecretsManagerLocation

The AWS Secrets Manager location configuration.

Contents

secretArn

The ARN of the AWS Secrets Manager secret containing the certificate.

Type: String

Pattern: `arn:aws(-[a-z-]+)?:secretsmanager:[a-z0-9-]+:[0-9]{12}:secret:[a-zA-Z0-9/_+=.@-]+`

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ServerDefinition

The MCP server definition with a schema version and inline content. The `schemaVersion` identifies the version of the MCP server configuration schema.

Contents

`inlineContent`

The inline content of the server definition.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 409600.

Required: No

`schemaVersion`

The schema version of the MCP server configuration. The schema version identifies the format of the server definition content.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

SessionFilter

Contains filter criteria for listing sessions.

Contents

eventFilter

The event filter condition to apply. Use this to filter sessions based on event presence.

Type: String

Valid Values: HAS_EVENTS

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

SessionFilterConfig

A time range filter for selecting sessions. Specifies the start and end times to narrow down which sessions are included.

Contents

endTime

The end time of the time range. Only sessions with activity before this timestamp are included.

Type: Timestamp

Required: No

startTime

The start time of the time range. Only sessions with activity at or after this timestamp are included.

Type: Timestamp

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

SessionLimits

The spending limits configuration for a payment session.

Contents

maxSpendAmount

The maximum amount that can be spent in the session.

Type: [Amount](#) object

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

SessionMetadataShape

Metadata for a specific session in a batch evaluation, including ground truth data and test scenario identifiers.

Contents

sessionId

The unique identifier of the session this metadata applies to.

Type: String

Required: Yes

groundTruth

The ground truth data for this session, including expected responses and assertions.

Type: [GroundTruthSource](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: No

metadata

Additional key-value metadata associated with this session.

Type: String to string map

Required: No

testScenarioId

An optional test scenario identifier for categorizing and tracking evaluation results.

Type: String

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

SessionSummary

Contains summary information about a session in an AgentCore Memory resource.

Contents

actorId

The identifier of the actor associated with the session.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[a-zA-Z0-9][a-zA-Z0-9-_/]*(?::[a-zA-Z0-9-_/]+)*[a-zA-Z0-9-_/]*`

Required: Yes

createdAt

The timestamp when the session was created.

Type: Timestamp

Required: Yes

sessionId

The unique identifier of the session.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 100.

Pattern: `[a-zA-Z0-9][a-zA-Z0-9-_]*`

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

SkillDefinition

The structured skill definition with a schema version and inline content.

Contents

inlineContent

The inline content of the skill definition.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 409600.

Required: No

schemaVersion

The schema version of the skill definition. If you don't specify a version, the service detects it automatically.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

SkillMdDefinition

The skill markdown definition for agent skills descriptors.

Contents

inlineContent

The inline markdown content of the skill definition.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 409600.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

SpanContext

The contextual information that uniquely identifies a span within the distributed tracing system. Contains session, trace, and span identifiers used to correlate evaluation results with specific agent execution points.

Contents

sessionId

The unique identifier of the session containing this span. Sessions represent complete conversation flows and are detected using configurable `SessionTimeoutMinutes` (default 15 minutes).

Type: String

Required: Yes

spanId

The unique identifier of the specific span being referenced. Spans represent individual operations like tool calls, model invocations, or other discrete actions within the agent's execution.

Type: String

Required: No

traceId

The unique identifier of the trace containing this span. Traces represent individual request-response interactions within a session and group related spans together.

Type: String

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

StreamUpdate

Contains information about an update to a stream.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

automationStreamUpdate

The update to an automation stream.

Type: [AutomationStreamUpdate](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

StripePrivyTokenRequestInput

Stripe Privy token request parameters.

Contents

requestBody

Request body JSON for the Privy API call.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 16384.

Pattern: `[\u0009\u000A\u000D\u0020-\u007E]+`

Required: Yes

requestPath

The path of the Stripe Privy API request.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: `/[a-zA-Z0-9/_\-\.\~%?=&]+`

Required: Yes

includeAuthorizationSignature

Set to true to generate privy-authorization-signature.

Type: Boolean

Required: No

requestHost

The host for the Privy API request. Defaults to "api.privy.io".

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: `[a-zA-Z0-9\-\.\]+`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

StripePrivyTokenResponseOutput

Stripe Privy token response containing `appId`, `basicAuthToken`, and optionally `authorizationSignature`.

Contents

`appId`

The Privy app ID for the `privy-app-id` header.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 512.

Pattern: `[a-zA-Z0-9\-_]+`

Required: Yes

`basicAuthToken`

Base64-encoded Basic Auth token (`appId:appSecret`) for the Authorization header.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 8192.

Required: Yes

`authorizationSignature`

Base64-encoded ECDSA P-256 authorization signature (only present when `includeAuthorizationSignature` is true).

Type: String

Length Constraints: Minimum length of 1. Maximum length of 8192.

Required: No

`requestExpiry`

Unix timestamp in milliseconds when the authorization signature expires.

Type: Long

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

SystemPromptConfig

The system prompt input, either as inline text or from a configuration bundle.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

configurationBundle

The system prompt sourced from a configuration bundle version.

Type: [SystemPromptConfigurationBundle](#) object

Required: No

text

The system prompt text provided inline.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 20000.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

SystemPromptConfigurationBundle

A system prompt sourced from a configuration bundle version.

Contents

bundleArn

The Amazon Resource Name (ARN) of the configuration bundle.

Type: String

Pattern: `arn:aws[a-zA-Z-]*:bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:configuration-bundle/[a-zA-Z][a-zA-Z0-9-_{0,99}]-[a-zA-Z0-9]{10}`

Required: Yes

systemPromptJsonPath

The JSON path within the configuration bundle that contains the system prompt.

Type: String

Required: Yes

versionId

The version identifier of the configuration bundle.

Type: String

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

SystemPromptRecommendationConfig

Configuration for generating system prompt optimization recommendations.

Contents

agentTraces

The agent traces to analyze for generating recommendations.

Type: [AgentTracesConfig](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: Yes

systemPrompt

The current system prompt to optimize.

Type: [SystemPromptConfig](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: Yes

evaluationConfig

The evaluation configuration specifying which evaluator to use for assessing recommendation quality.

Type: [RecommendationEvaluationConfig](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)

- [AWS SDK for Ruby V3](#)

SystemPromptRecommendationResult

The result of a system prompt recommendation, containing the optimized prompt.

Contents

configurationBundle

The configuration bundle containing the recommended system prompt, if the input was sourced from a configuration bundle.

Type: [RecommendationResultConfigurationBundle](#) object

Required: No

errorCode

The error code if the recommendation failed.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 1024.

Required: No

errorMessage

The error message if the recommendation failed.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Required: No

explanation

An explanation of why the recommendation was generated and what patterns were identified in the agent traces.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 4096.

Required: No

recommendedSystemPrompt

The optimized system prompt text generated by the recommendation.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 20000.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

TargetRef

A reference to a gateway target.

Contents

name

The name of the gateway target.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 100.

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

TokenBalance

A single token balance entry.

Contents

amount

Raw balance in the smallest denomination (e.g., USDC base units where 1 USDC = 1000000).

Type: String

Required: Yes

chain

The specific blockchain chain.

Type: String

Valid Values: BASE | BASE_SEPOLIA | ETHEREUM | SOLANA | SOLANA_DEVNET

Required: Yes

decimals

Number of decimal places for the token (e.g., 6 for USDC).

Type: Integer

Required: Yes

network

The blockchain network family (ETHEREUM or SOLANA).

Type: String

Valid Values: ETHEREUM | SOLANA

Required: Yes

token

The supported token for this balance.

Type: String

Valid Values: USDC

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

TokenUsage

The token consumption statistics for language model operations during evaluation. Provides detailed breakdown of input, output, and total tokens used for cost tracking and performance monitoring.

Contents

inputTokens

The number of tokens consumed for input processing during the evaluation. Includes tokens from the evaluation prompt, agent traces, and any additional context provided to the evaluator model.

Type: Integer

Required: No

outputTokens

The number of tokens generated by the evaluator model in its response. Includes tokens for the score, explanation, and any additional output produced during the evaluation process.

Type: Integer

Required: No

totalTokens

The total number of tokens consumed during the evaluation, calculated as the sum of input and output tokens. Used for cost calculation and rate limiting within the service limits.

Type: Integer

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)

- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ToolArguments

The collection of arguments that specify the operation to perform and its parameters when invoking a tool in Amazon Bedrock AgentCore. Different tools require different arguments, and this structure provides a flexible way to pass the appropriate arguments to each tool type.

Contents

clearContext

Whether to clear the context for the tool.

Type: Boolean

Required: No

code

The code to execute in a code interpreter session. This is the source code in the specified programming language that will be executed by the code interpreter.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 100000000.

Required: No

command

The command to execute with the tool.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 100000000.

Required: No

content

The content for the tool operation.

Type: Array of [InputContentBlock](#) objects

Required: No

directoryPath

The directory path for the tool operation.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 100000000.

Required: No

language

The programming language of the code to execute. This tells the code interpreter which language runtime to use for execution.

Type: String

Valid Values: python | javascript | typescript

Required: No

path

The path for the tool operation.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 100000000.

Required: No

paths

The paths for the tool operation.

Type: Array of strings

Length Constraints: Minimum length of 0. Maximum length of 100000000.

Required: No

runtime

The runtime environment to use for code execution. If not specified, defaults to deno for JavaScript and TypeScript.

Type: String

Valid Values: nodejs | deno | python

Required: No

taskId

The identifier of the task for the tool operation.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 100000000.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ToolDescriptionConfig

The tool description content.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

text

The tool description as inline text.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 20000.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ToolDescriptionConfigurationBundle

Tool descriptions sourced from a configuration bundle version.

Contents

bundleArn

The Amazon Resource Name (ARN) of the configuration bundle.

Type: String

Pattern: `arn:aws[a-zA-Z-]*:bedrock-agentcore:[a-z0-9-]+:[0-9]{12}:configuration-bundle/[a-zA-Z][a-zA-Z0-9-_{0,99}]-[a-zA-Z0-9]{10}`

Required: Yes

tools

The list of tool entries mapping tool names to their JSON paths within the bundle.

Type: Array of [ConfigurationBundleToolEntry](#) objects

Required: Yes

versionId

The version identifier of the configuration bundle.

Type: String

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ToolDescriptionInput

A tool description input containing the tool name and its current description.

Contents

toolDescription

The current description of the tool to optimize.

Type: [ToolDescriptionConfig](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: Yes

toolName

The name of the tool.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: `[a-zA-Z0-9_-\.\.]+`

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ToolDescriptionOutput

The output for a single tool description recommendation.

Contents

toolName

The name of the tool.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: `[a-zA-Z0-9_-\.\.]+`

Required: Yes

explanation

An explanation of why the recommendation was generated for this tool and what patterns were identified in the agent traces.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 4096.

Required: No

recommendedToolDescription

The optimized tool description text generated by the recommendation.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 20000.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ToolDescriptionRecommendationConfig

Configuration for generating tool description optimization recommendations.

Contents

agentTraces

The agent traces to analyze for generating tool description recommendations.

Type: [AgentTracesConfig](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: Yes

toolDescription

The current tool descriptions to optimize.

Type: [ToolDescriptionSource](#) object

Note: This object is a Union. Only one member of this object can be specified or returned.

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ToolDescriptionRecommendationResult

The result of a tool description recommendation, containing optimized descriptions.

Contents

configurationBundle

The configuration bundle containing the recommended tool descriptions, if the input was sourced from a configuration bundle.

Type: [RecommendationResultConfigurationBundle](#) object

Required: No

errorCode

The error code if the recommendation failed.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 1024.

Required: No

errorMessage

The error message if the recommendation failed.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Required: No

tools

The list of tools with their recommended descriptions.

Type: Array of [ToolDescriptionOutput](#) objects

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ToolDescriptionSource

The source of tool descriptions, either inline text or from a configuration bundle.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

configurationBundle

Tool descriptions sourced from a configuration bundle version.

Type: [ToolDescriptionConfigurationBundle](#) object

Required: No

toolDescriptionText

Tool descriptions provided as inline text.

Type: [ToolDescriptionTextInput](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ToolDescriptionTextInput

Inline tool description input containing a list of tools.

Contents

tools

The list of tool descriptions to optimize.

Type: Array of [ToolDescriptionInput](#) objects

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ToolResultStructuredContent

Contains structured content from a tool result.

Contents

executionTime

The execution time of the tool operation in milliseconds.

Type: Double

Required: No

exitCode

The exit code from the tool execution.

Type: Integer

Required: No

stderr

The standard error output from the tool execution.

Type: String

Required: No

stdout

The standard output from the tool execution.

Type: String

Required: No

taskId

The identifier of the task that produced the result.

Type: String

Required: No

taskStatus

The status of the task that produced the result.

Type: String

Valid Values: submitted | working | completed | canceled | failed

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ToolsDefinition

The MCP tools definition with a protocol version and inline content. The `protocolVersion` identifies the MCP protocol version that the tools conform to. This differs from `schemaVersion` in the server definition, which identifies the server configuration schema format.

Contents

inlineContent

The inline content of the tools definition.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 409600.

Required: No

protocolVersion

The MCP protocol version that the tools conform to. This differs from the `schemaVersion` field in the server definition, which identifies the server configuration schema format.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

UserIdentifier

The OAuth2.0 token or user ID that was used to generate the workload access token used for initiating the user authorization flow to retrieve OAuth2.0 tokens.

Contents

Important

This data type is a UNION, so only one of the following members can be specified when used or returned.

userId

The ID of the user for whom you have retrieved a workload access token for

Type: String

Length Constraints: Minimum length of 1. Maximum length of 128.

Required: No

userToken

The OAuth2.0 token issued by the user's identity provider that was used to generate the workload access token

Type: String

Length Constraints: Minimum length of 1. Maximum length of 131072.

Pattern: `[A-Za-z0-9-_=]+.[A-Za-z0-9-_=]+.[A-Za-z0-9-_=]+`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

UserIntentAffectedSession

A session associated with a user intent cluster.

Contents

sessionId

The unique identifier of the session.

Type: String

Required: Yes

userMessages

The user messages from this session that contributed to the intent cluster.

Type: Array of strings

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

UserIntentCluster

A cluster of similar user intents identified across sessions.

Contents

affectedSessionCount

The number of sessions with this user intent.

Type: Integer

Required: Yes

affectedSessions

The list of sessions with this user intent.

Type: Array of [UserIntentAffectedSession](#) objects

Array Members: Minimum number of 0 items.

Required: Yes

clusterId

The unique identifier of the user intent cluster.

Type: Integer

Required: Yes

description

A description of the user intent pattern.

Type: String

Required: Yes

name

The name of the user intent cluster.

Type: String

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

UserIntentClusteringResultContent

The user intent clustering result containing grouped user intents identified across evaluated sessions.

Contents

userIntents

The list of user intent clusters identified across analyzed sessions.

Type: Array of [UserIntentCluster](#) objects

Array Members: Minimum number of 0 items.

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ValidationExceptionField

Stores information about a field passed inside a request that resulted in an exception.

Contents

message

A message describing why this field failed validation.

Type: String

Required: Yes

name

The name of the field.

Type: String

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Variant

A variant in an A/B test, representing either the control (C) or treatment (T1) configuration.

Contents

name

The name of the variant. Must be C for control or T1 for treatment.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2.

Pattern: (C|T1)

Required: Yes

variantConfiguration

The configuration for this variant, including the configuration bundle or target reference.

Type: [VariantConfiguration](#) object

Required: Yes

weight

The percentage of traffic to route to this variant. Weights across all variants must sum to 100.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)

- [AWS SDK for Ruby V3](#)

VariantConfiguration

The configuration for an A/B test variant.

Contents

configurationBundle

A reference to a configuration bundle version to use for this variant.

Type: [ConfigurationBundleRef](#) object

Required: No

target

A reference to a gateway target to route traffic to for this variant.

Type: [TargetRef](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

VariantResult

Statistical results for a treatment variant compared against the control.

Contents

isSignificant

Whether the observed difference is statistically significant.

Type: Boolean

Required: Yes

mean

The mean evaluation score for this variant.

Type: Double

Required: Yes

sampleSize

The number of sessions evaluated for this variant.

Type: Integer

Required: Yes

variantName

The name of the treatment variant.

Type: String

Required: Yes

absoluteChange

The absolute change in mean score compared to the control variant.

Type: Double

Required: No

confidenceInterval

The confidence interval for the observed difference.

Type: [ConfidenceInterval](#) object

Required: No

percentChange

The percentage change in mean score compared to the control variant.

Type: Double

Required: No

pValue

The p-value indicating the statistical significance of the observed difference.

Type: Double

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ViewPort

The configuration that defines the dimensions of a browser viewport in a browser session. The viewport determines the visible area of web content and affects how web pages are rendered and displayed. Proper viewport configuration ensures that web content is displayed correctly for the agent's browsing tasks.

Contents

height

The height of the viewport in pixels. This value determines the vertical dimension of the visible area. Valid values range from 600 to 1080 pixels.

Type: Integer

Valid Range: Minimum value of 240. Maximum value of 2160.

Required: Yes

width

The width of the viewport in pixels. This value determines the horizontal dimension of the visible area. Valid values range from 800 to 1920 pixels.

Type: Integer

Valid Range: Minimum value of 320. Maximum value of 3840.

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Common Parameters

The following list contains the parameters that all actions use for signing Signature Version 4 requests with a query string. Any action-specific parameters are listed in the topic for that action. For more information about Signature Version 4, see [Signing AWS API requests](#) in the *IAM User Guide*.

X-Amz-Algorithm

The hash algorithm that you used to create the request signature.

Condition: Specify this parameter when you include authentication information in a query string instead of in the HTTP authorization header.

Type: string

Valid Values: AWS4-HMAC-SHA256

Required: Conditional

X-Amz-Credential

The credential scope value, which is a string that includes your access key, the date, the region you are targeting, the service you are requesting, and a termination string ("aws4_request"). The value is expressed in the following format: *access_key/YYYYMMDD/region/service/aws4_request*.

For more information, see [Create a signed AWS API request](#) in the *IAM User Guide*.

Condition: Specify this parameter when you include authentication information in a query string instead of in the HTTP authorization header.

Type: string

Required: Conditional

X-Amz-Date

The date that is used to create the signature. The format must be ISO 8601 basic format (YYYYMMDD'T'HHMMSS'Z'). For example, the following date time is a valid X-Amz-Date value: 20120325T120000Z.

Condition: X-Amz-Date is optional for all requests; it can be used to override the date used for signing requests. If the Date header is specified in the ISO 8601 basic format, X-Amz-Date is not required. When X-Amz-Date is used, it always overrides the value of the Date header. For more information, see [Elements of an AWS API request signature](#) in the *IAM User Guide*.

Type: string

Required: Conditional

X-Amz-Security-Token

The temporary security token that was obtained through a call to AWS Security Token Service (AWS STS). For a list of services that support temporary security credentials from AWS STS, see [AWS services that work with IAM](#) in the *IAM User Guide*.

Condition: If you're using temporary security credentials from AWS STS, you must include the security token.

Type: string

Required: Conditional

X-Amz-Signature

Specifies the hex-encoded signature that was calculated from the string to sign and the derived signing key.

Condition: Specify this parameter when you include authentication information in a query string instead of in the HTTP authorization header.

Type: string

Required: Conditional

X-Amz-SignedHeaders

Specifies all the HTTP headers that were included as part of the canonical request. For more information about specifying signed headers, see [Create a signed AWS API request](#) in the *IAM User Guide*.

Condition: Specify this parameter when you include authentication information in a query string instead of in the HTTP authorization header.

Type: string

Required: Conditional

Common Error Types

This section lists common error types that this AWS service may return. Not all services return all error types listed here. For errors specific to an API action for this service, see the topic for that API action.

AccessDeniedException

You don't have permission to perform this action. Verify that your IAM policy includes the required permissions.

HTTP Status Code: 403

ExpiredTokenException

The security token included in the request has expired. Request a new security token and try again.

HTTP Status Code: 403

IncompleteSignature

The request signature doesn't conform to AWS standards. Verify that you're using valid AWS credentials and that your request is properly formatted. If you're using an SDK, ensure it's up to date.

HTTP Status Code: 403

InternalFailure

The request can't be processed right now because of an internal server issue. Try again later. If the problem persists, contact AWS Support.

HTTP Status Code: 500

MalformedHttpRequestException

The request body can't be processed. This typically happens when the request body can't be decompressed using the specified content encoding algorithm. Verify that the content encoding header matches the compression format used.

HTTP Status Code: 400

NotAuthorized

You don't have permissions to perform this action. Verify that your IAM policy includes the required permissions.

HTTP Status Code: 401

OptInRequired

Your AWS account needs a subscription for this service. Verify that you've enabled the service in your account.

HTTP Status Code: 403

RequestAbortedException

The request was aborted before a response could be returned. This typically happens when the client closes the connection.

HTTP Status Code: 400

RequestEntityTooLargeException

The request entity is too large. Reduce the size of the request body and try again.

HTTP Status Code: 413

RequestTimeoutException

The request timed out. The server didn't receive the complete request within the expected time frame. Try again.

HTTP Status Code: 408

ServiceUnavailable

The service is temporarily unavailable. Try again later.

HTTP Status Code: 503

ThrottlingException

Your request rate is too high. The AWS SDKs automatically retry requests that receive this exception. Reduce the frequency of requests.

HTTP Status Code: 400

UnknownOperationException

The action or operation isn't recognized. Verify that the action name is spelled correctly and that it's supported by the API version you're using.

HTTP Status Code: 404

UnrecognizedClientException

The X.509 certificate or AWS access key ID you provided doesn't exist in our records. Verify that you're using valid credentials and that they haven't expired.

HTTP Status Code: 403

ValidationError

The input doesn't meet the required format or constraints. Check that all required parameters are included and that values are valid.

HTTP Status Code: 400