



Operacionalización de la IA de los agentes en AWS

AWS Orientación prescriptiva



AWS Orientación prescriptiva: Operacionalización de la IA de los agentes en AWS

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

Las marcas comerciales y la imagen comercial de Amazon no se pueden utilizar en relación con ningún producto o servicio que no sea de Amazon, de ninguna manera que pueda causar confusión entre los clientes y que menosprecie o desacredite a Amazon. Todas las demás marcas registradas que no son propiedad de Amazon son propiedad de sus respectivos propietarios, que pueden o no estar afiliados, conectados o patrocinados por Amazon.

Table of Contents

Introducción	1
Áreas de interés	1
Destinatarios previstos	2
Objetivos	2
Acerca de esta serie de contenido	3
Fundamentos de la IA agencial	4
Áreas de interés	6
Intención y alcance	7
Strategy (Estrategia)	7
Valor empresarial	9
Composabilidad y colaboración	10
Strategy (Estrategia)	10
Valor empresarial	13
Tenencia múltiple y control	13
Strategy (Estrategia)	14
Valor empresarial	14
Autonomía confiable	15
Strategy (Estrategia)	16
Valor empresarial	17
Administración del ciclo de vida	18
Strategy (Estrategia)	18
Valor empresarial	19
Alineación empresarial	20
Strategy (Estrategia)	20
Entrega de software	22
Zonas de intención	22
Evolución del SDLC	23
Preparar equipos	25
Preparándose para la expansión	26
Equipos y modelos de propiedad	26
Administración de cambios	27
Interoperabilidad y colaboración	28
Gobernanza	29
Mentalidad operativa	30

Escalado	30
Conclusión	32
Recursos	34
Servicios de AWS	34
Otros recursos AWS	35
Historial de documentos	37
Glosario	38
#	38
A	39
B	42
C	44
D	48
E	52
F	54
G	56
H	57
I	59
L	61
M	63
O	67
P	70
Q	73
R	73
S	76
T	80
U	82
V	83
W	83
Z	84
.....	lxxxvi

Operacionalización de la IA de los agentes en AWS

Aaron Sempff, Brad Ryan, Bhargs Srivathsan y Akhil Bhaskar, de Amazon Web Services

Agosto [de](#) 2025 (historial del documento)

La IA de los agentes no es una característica, es un nuevo paradigma operativo. Organismos que inviertan en una arquitectura disciplinada, marcos de confianza y modelos de implementación alineados con los negocios liderarán la próxima generación de empresas inteligentes y adaptables.

La IA de agencia representa la convergencia de los agentes de software autónomos y la IA generativa. Fusiona la toma de decisiones y el comportamiento orientado a objetivos de los agentes con las capacidades de comprensión y generación del lenguaje de los grandes modelos lingüísticos (LLMs). Estos agentes pueden razonar, actuar, adaptarse y colaborar en entornos empresariales dinámicos. Para aprovechar este potencial, las empresas deben cambiar su forma de pensar y pasar de la implementación de modelos a la infraestructura de agentes.

Esta guía proporciona una estrategia organizacional para transformar la IA de los agentes, pasando de ser experimentos aislados a convertirse en una infraestructura a escala empresarial que genere valor. Puede ayudarlo a integrar agentes inteligentes en todos los flujos de trabajo con gobernanza, escalabilidad y alineación empresarial.

Áreas de enfoque y recomendaciones clave

Esta guía se centra en las siguientes áreas fundamentales a la hora de poner en práctica la IA de los agentes. Se proporcionan recomendaciones organizativas y empresariales para cada área de enfoque:

- [Área de enfoque 1: Aclarar la intención y el alcance del agente](#)— Alinee a los agentes con las prioridades empresariales y los cuellos de botella cognitivos. Trate a los agentes como compañeros de equipo digitales, no solo como herramientas.
- [Área de enfoque 2: Diseño para la componibilidad y la colaboración](#)— Adopte sistemas multiagente con una arquitectura modular, protocolos semánticos y delegación dinámica a través de agentes árbitros.
- [Área de enfoque 3: Diseñar soluciones para el control y la multitenencia](#)— Cree una infraestructura escalable y adaptada a los inquilinos con servicios de agentes compartidos, una gobernanza centralizada y un acceso basado en funciones.

- [Área de enfoque 4: Generar confianza a través de la identidad, las barreras y la observabilidad](#)— Imponga la trazabilidad, los controles de tiempo de ejecución y la explicabilidad para ganarse la confianza de las partes interesadas.
- [Área de enfoque 5: Gestionar el ciclo de vida](#)— Establezca procesos continuos de integración y despliegue continuo (CI/CD), un rápido control de versiones, telemetría y reentrenamiento continuo para respaldar el rendimiento y la eficiencia de la IA de los agentes.
- [Área de enfoque 6: Alinee los modelos de los agentes con los modelos de negocio](#)— Monetice las capacidades de los agentes mediante modelos basados en el uso, métricas internas de ROI y ofertas comerciales.

Puede utilizar las recomendaciones de esta guía para preparar su empresa para la IA de los agentes a gran escala. Describe cómo las organizaciones deben reestructurarse en torno a la IA de los agentes, lo que incluye crear equipos DevOps para agentes (AgentOps), sistemas interoperables y estrategias de gestión del cambio que escalen la adopción. Hace hincapié en pensar primero en la toma de decisiones y en la alineación con el Marco de AWS Buena Arquitectura.

Destinatarios previstos

Esta guía está dirigida a arquitectos empresariales, líderes de AI/ML ingeniería y estrategias de transformación digital que están diseñando y escalando sistemas de agencias, integrando la IA en los flujos de trabajo empresariales principales y LLMs operacionalizando agentes autónomos en entornos de producción. Para comprender los conceptos y las recomendaciones de esta guía, debe estar familiarizado con las arquitecturas modernas nativas de la nube y los sistemas distribuidos, los grandes modelos de lenguaje, las capacidades de los modelos básicos y los principios de la gobernanza de la IA y la ingeniería de plataformas. DevOps

Objetivos

Al implementar las recomendaciones de esta guía, su organización puede lograr los siguientes resultados empresariales:

- Acelera la toma de decisiones y la ejecución del flujo de trabajo a través de agentes autónomos y orientados a los objetivos que reducen los cuellos de botella humanos y la carga cognitiva.
- Implementaciones escalables y rentables de capacidades inteligentes en todas las unidades de negocio, a través de plataformas de agentes multiusuario reutilizables.

- Mayor resiliencia, confianza y gobernanza en los sistemas de IA, lo que permite una adopción segura en entornos regulados, de misión crítica o orientados al cliente.

Acerca de esta serie de contenido

Esta guía forma parte de una serie sobre la IA de los agentes en AWS. Para obtener más información y ver las demás guías de esta serie, consulte [Agentic AI](#) en el sitio web de orientación prescriptiva. AWS

Fundamentos estratégicos de la IA agencial

Los sistemas de agencia no son nuevos. Los agentes de software, incluida la automatización robótica de procesos (RPA) y los motores de decisión, existen desde hace décadas. Pero eran simples y deterministas, y estaban diseñados para seguir reglas predefinidas y una lógica simbólica para ejecutar tareas repetitivas y de baja variación. Con el auge de la IA generativa, las reglas del juego han cambiado. Los modelos lingüísticos de gran tamaño (LLMs) ahora pueden interpretar entradas complejas, generar respuestas de forma dinámica y sintetizar rápidamente el conocimiento. Ahora puede escalar la agencia sin una lógica frágil o codificada. Ahora, los agentes pueden razonar, tomar decisiones, utilizar herramientas, adaptarse al contexto y coordinarse con otros agentes en todos los flujos de trabajo. Pueden operar de forma autónoma para alcanzar los objetivos, mantener la memoria y reflexionar sobre los resultados.

Sin embargo, la capacidad bruta no es suficiente. La inteligencia sin integración produce novedad, no impacto. Para sacar el máximo partido a los poderosos LLMs, las empresas deben pasar de los experimentos aislados a los ecosistemas diseñados. Los agentes deben ser tratados como servicios de nivel de producción que operan bajo la misma disciplina que cualquier sistema empresarial. Esto incluye la gobernanza, la observabilidad, los modelos de identidad seguros y la gestión del ciclo de vida. También deben generar resultados comerciales reales, no un potencial especulativo. Estos sistemas deben diseñarse con límites claros para la toma de decisiones y tolerancia a los fallos. Es importante incorporar mecanismos de recuperación automatizados, monitoreo del rendimiento en tiempo real y administración de recursos escalable. Esto le ayuda a gestionar la naturaleza dinámica y no determinista de las interacciones entre los agentes y, al mismo tiempo, a mantener niveles de servicio uniformes en todos los flujos de trabajo empresariales.

A nivel fundamental, las empresas deben replantearse la forma en que la inteligencia está integrada en la estructura de las operaciones. Los agentes deben estar diseñados para integrarse con los sistemas principales, cumplir con las políticas empresariales y ofrecer un valor cuantificable. Deben operar a escala, en todos los departamentos, dominios y contextos de usuario. La puesta en práctica de la IA de los agentes se basa, en última instancia, en el uso; es la diferencia entre implementar una IA que realice tareas aisladas y desplegar agentes que hagan evolucionar su modelo de negocio.

La IA de Agentic representa una nueva filosofía operativa que requiere un cambio fundamental en la forma en que abordamos los sistemas, los procesos y las personas para ampliar la inteligencia en toda la organización. Los agentes se convierten en activos estratégicos que amplifican las capacidades humanas. Al integrar la IA de los agentes en sus operaciones, las organizaciones

pueden obtener información valiosa que impulsa el valor empresarial, aumenta las capacidades humanas y optimiza los flujos de trabajo complejos.

Áreas de enfoque estratégico para la IA de los agentes

Para pasar de los primeros prototipos a sistemas aptos para la producción y generadores de valor, los equipos necesitan una estrategia coherente que combine la arquitectura, los procesos y la concepción del producto.

Muchas organizaciones siguen abordando la IA con una mentalidad centrada en las herramientas o en los modelos. La IA generativa ha amplificado la experimentación, pero a menudo no se alinea claramente con la estrategia empresarial ni con los resultados mensurables. Sin una función estratégica definida, los agentes corren el riesgo de convertirse en experimentos novedosos que agotan los recursos en lugar de ofrecer un valor escalable. Para establecer el papel estratégico de la IA de los agentes, las organizaciones deben comenzar con las prioridades empresariales. Identifique las áreas de sobrecarga cognitiva, los cuellos de botella en la toma de decisiones o los flujos de trabajo fragmentados en los que la autonomía puede aliviar. Utilice enunciados de problemas específicos para determinar las responsabilidades de los agentes. Trate a los agentes como compañeros de equipo digitales, no como herramientas, que pueden razonar, delegar y adaptarse.

Las ciencias de la decisión son la disciplina que combina la ciencia de datos, el análisis y el modelado del comportamiento para mejorar la toma de decisiones. Debe integrarse en una fase temprana del proceso de arquitectura de los agentes para alinear el diseño con los resultados empresariales. Al identificar los patrones de decisión, simular las compensaciones y cuantificar el impacto en el valor, las ciencias de la toma de decisiones pueden ayudarlo a determinar dónde la autonomía de los agentes puede ofrecer el mayor valor. Las ciencias de la toma de decisiones pueden acelerar las decisiones, reducir los errores y permitir adaptaciones en tiempo real. Esta base basada en datos basa el diseño de los agentes en información mensurable y permite una mayor integración con las tecnologías empresariales existentes, como los motores de reglas, las plataformas de análisis y los modelos predictivos.

Para ayudar a establecer el papel estratégico de los agentes, en esta sección se presentan las áreas de enfoque fundamentales que constituyen la columna vertebral de la operacionalización de la IA de los agentes. Cada una de ellas se basa en un trabajo fundamental que debe realizarse desde la perspectiva de un líder técnico, arquitecto o propietario de un producto, que es responsable de la concepción y el diseño de los agentes. Estas áreas de enfoque no son pasos secuenciales. Vale la pena revisar cada una de ellas a lo largo del ciclo de vida del sistema para desarrollar ecosistemas de agentes resilientes, escalables y monetizables.

Esta sección contiene las siguientes áreas de enfoque:

- [Área de enfoque 1: Aclarar la intención y el alcance del agente](#)
- [Área de enfoque 2: Diseño para la componibilidad y la colaboración](#)
- [Área de enfoque 3: Diseñar soluciones para el control y la multitenencia](#)
- [Área de enfoque 4: Generar confianza a través de la identidad, las barreras y la observabilidad](#)
- [Área de enfoque 5: Gestionar el ciclo de vida](#)
- [Área de enfoque 6: Alinee los modelos de los agentes con los modelos de negocio](#)

Área de enfoque 1: Aclarar la intención y el alcance del agente

Trabajo por hacer: «Ayúdame a asegurarme de que cada agente resuelva un problema real con límites claros, no solo con una buena demostración».

La IA de los agentes no se limita a desarrollar capacidades. Se trata de resolver el problema correcto, de la manera correcta y con el resultado correcto. Para empezar, hay que tener completamente claro el propósito de la solución de IA de la agencia.

Strategy (Estrategia)

Con demasiada frecuencia, las organizaciones comienzan con lo que el modelo puede hacer (por ejemplo APIs, llamar, responder preguntas o generar resúmenes) y adaptan un caso de uso en función de ello. Esto se traduce en un aumento del alcance, una mala integración y unos agentes que son técnicamente impresionantes pero inútiles desde el punto de vista operativo. En su lugar, comience por definir el papel del agente mediante preguntas específicas como las siguientes:

- ¿De qué resultado específico es responsable el agente?
- ¿En nombre de quién actúa?
- ¿Quién se beneficia?
- ¿Dónde comienza y termina la autonomía del agente?
- ¿Qué pasa cuando falla?

Un agente bien definido tiene un trabajo claro, responsabilidades definidas y criterios de éxito cuantificables. No pienses en el agente como un asistente o un chatbot. En vez de eso, asígnele un cargo. Considéralo un agente de atención al cliente, un gestor de devoluciones de productos o un supervisor de cumplimiento.

Cuando interactúes con las partes interesadas o los clientes, haz hincapié en la escalabilidad y adaptabilidad de los sistemas de inteligencia artificial de los agentes. Estos agentes evolucionan con la empresa y mejoran continuamente a través del aprendizaje y la retroalimentación. Para reducir la resistencia y acelerar la adopción, destaque cómo las herramientas de los agentes se diseñan teniendo en cuenta la empatía de los trabajadores. Proporcionan transparencia, control y mecanismos de anulación opcionales que generan confianza. En lugar de sustituir a las personas, los agentes aumentan la capacidad humana y la toma de decisiones, lo que ayuda a los empleados a mantenerse informados y a centrarse en tareas de gran valor.

La clave para una implementación exitosa es alinear la IA de las agencias con resultados empresariales específicos y de alto impacto. Anime a los equipos y socios a comenzar con proyectos piloto específicos que resuelvan los puntos débiles visibles. Las ganancias rápidas generan un retorno de la inversión (ROI) mensurable, fomentan la aceptación interna y generan el impulso necesario para una adopción más amplia.

Para guiar la adopción y la madurez, las organizaciones pueden encuadrar el diseño de los agentes en un modelo evolutivo. La autonomía, la complejidad y el impacto empresarial de los agentes aumentan progresivamente. Las etapas de este modelo son las siguientes:

- Los agentes observadores obtienen información a partir del ruido. Un ejemplo es un agente del sentimiento del mercado que rastrea la percepción de la marca en los canales digitales.
- Los agentes asistentes apoyan la toma de decisiones humanas. Un ejemplo es un agente asesor de negocios que sintetiza los datos de la competencia y las condiciones del mercado para los equipos de ventas.
- Los agentes autónomos actúan de forma independiente dentro de unos límites definidos. Un ejemplo es un agente de asignación de recursos que ajusta dinámicamente la infraestructura de la nube en función de la demanda.
- Los agentes de Orchestrator coordinan los flujos de trabajo de varios agentes. Un ejemplo es un agente de optimización de la cadena de suministro que gestiona las interacciones entre los agentes de inventario, logística y previsión.
- Los agentes innovadores generan nuevas posibilidades estratégicas. Un ejemplo es un agente de innovación de modelos de negocio que analiza las tendencias del mercado y recomienda nuevas fuentes de ingresos.

Enmarcar a los agentes en torno a estos resultados estratégicos y niveles de madurez aumenta la concentración, acelera la adopción y fomenta la confianza de las partes interesadas.

Para respaldar la alineación en esta área de enfoque Servicios de AWS, como [Amazon Quick](#), puede visualizar los indicadores clave de rendimiento (KPIs) que están vinculados a los resultados impulsados por los agentes. Puedes usar [Amazon CloudWatch](#) para monitorear el comportamiento de los agentes, las métricas de rendimiento y el estado del sistema prácticamente en tiempo real. Utilice los comentarios operativos para ajustar las interacciones de los agentes y el uso de los recursos. [AWS CloudTrail](#) puede proporcionar visibilidad de la actividad de los agentes y los patrones de integración durante las primeras fases de experimentación y perfeccionamiento.

El valor empresarial de definir la intención y el alcance

La adopción de la IA institucional representa un cambio fundamental en la forma en que las organizaciones abordan la transformación digital y la excelencia operativa. No se trata únicamente de la automatización. Se trata de permitir una autonomía inteligente que acelere la toma de decisiones y la obtención de valores.

Entre los principales impulsores empresariales se incluyen los siguientes:

- **Ventaja competitiva:** los primeros en adoptarlo obtienen una ventaja estratégica al obtener información más rápida, un mejor servicio y operaciones adaptables.
- **Mejora de la experiencia del cliente:** los agentes ofrecen asistencia en tiempo real, personalizada y siempre disponible que aumenta la satisfacción y la lealtad.
- **Eficiencia operativa:** la IA de Agentic reduce significativamente la carga cognitiva humana al automatizar tareas de toma de decisiones complejas y repetitivas. Esto permite al personal centrarse en actividades de mayor valor y puede reducir los costes.

Los casos de uso reales en todos los sectores incluyen los siguientes:

- **Servicios financieros:** los agentes de IA podrían ofrecer asesoramiento financiero personalizado y detectar el fraude.
- **Atención médica:** los agentes de los planes de tratamiento y clasificación podrían mejorar el rendimiento clínico.
- **Venta minorista:** los agentes pueden actuar como asistentes de compras inteligentes u optimizar el inventario en tiempo real.
- **Fabricación:** los agentes pueden realizar un mantenimiento predictivo o coordinar las cadenas de suministro.

Área de enfoque 2: Diseño para la componibilidad y la colaboración

Trabajo pendiente: «Permítanme crear agentes como construyo servicios: modulares y comprobables, de modo que se puedan componer y organizar según sea necesario».

Muchas iniciativas de IA comienzan como proyectos piloto monolíticos y centrados en modelos. Son útiles, pero es difícil escalarlos entre dominios o adaptarse a problemas complejos. Valora los compuestos cuando estos agentes están diseñados para interoperar. En tecnología, la componibilidad es el acto de combinar componentes modulares para crear una solución flexible y escalable que pueda adaptarse a los cambios. Sin la componibilidad, la inteligencia queda encerrada en flujos de trabajo específicos. Además, la colaboración entre agentes introduce complejidades de organización, administración del estado y negociación de protocolos que los equipos de automatización tradicionales podrían no estar preparados para gestionar.

Strategy (Estrategia)

Adopte el paradigma de los múltiples agentes. Modele agentes como departamentos organizacionales: modulares, especializados e interoperables. Defina interfaces claras, formatos de contexto compartidos y protocolos de comunicación estándar, como [Model Context Protocol \(MCP\)](#) o [Agent2Agent \(A2A\)](#). Adopte patrones de orquestación multiagente, como la coordinación jerárquica, mediante enjambres o grafos. Estos patrones ayudan a los agentes a descubrir capacidades y solicitar servicios entre sí de forma dinámica, ya sea en flujos de trabajo paralelos, secuenciales o basados en el consenso, según la estructura de la tarea y el nivel de confianza.

Para promover una colaboración escalable y gobernada, utilice un agente árbitro. Este tipo de agente es una autoridad neutral que facilita la delegación de tareas en función de las capacidades conocidas y las estrategias alternativas. Si bien no es un controlador centralizado, un agente árbitro desempeña un papel fundamental en materia de confianza y cumplimiento. Garantiza que las tareas delicadas o reguladas se envíen únicamente a los agentes que cumplan con los requisitos de identidad y políticas. Actúa como un guardián de los flujos de trabajo sujetos a políticas. Refuerza el aislamiento y permite una delegación explicable. Fundamentalmente, un agente árbitro no es un cuello de botella, sino que coexiste con agentes que se autocoordinan y actúan de manera horizontal. peer-to-peer Estos agentes delegan subtareas, comparten el contexto y resuelven las dependencias directamente.

Este modelo híbrido admite tanto la asignación determinista (a través del agente árbitro) como la colaboración emergente. Combina estructura con flexibilidad. Dentro de esta arquitectura, los agentes se pueden clasificar en las siguientes funciones especializadas:

- Agentes de toma de decisiones, como los encargados de hacer cumplir las políticas, los asignadores de recursos y los evaluadores de riesgos
- Agentes de conocimiento, como agregadores de contexto, reconocedores de patrones y detectores de anomalías
- Agentes de ejecución, como ejecutores de tareas, controladores de calidad y gerentes de integración

Para coordinarse de forma eficaz, los sistemas multiagente deben admitir protocolos de interacción sólidos para la gestión del estado, la recuperación de fallos y la resolución de conflictos. Esto promueve la estabilidad y la responsabilidad, incluso cuando los agentes operan de forma independiente.

Establezca reglas claras para el escalado, como la instanciación de los agentes en función de la carga, la asignación de recursos en función del contexto y la detección y el registro automatizados de capacidades. Estas medidas ayudan al sistema a crecer de forma dinámica en respuesta a la demanda o la complejidad.

Diseñe los agentes para que sean ready-to-use módulos dentro de un sustrato de mensajería distribuido. Por ejemplo, puedes usar [Amazon EventBridge](#) con A2A o MCP en lugar de servicios aislados. Adopte plantillas de versiones, CI/CD canalizaciones y agentes para respaldar la estabilidad del sistema y, al mismo tiempo, acelerar la adopción interna y la evolución del ciclo de vida. Fomente la reutilización y la estandarización del código para reducir la fricción de la integración y promover un ecosistema resiliente.

La colaboración es un factor multiplicador de fuerzas. Facilita la escalabilidad, la especialización y la resiliencia en entornos con varios agentes. Para respaldar esta colaboración dinámica, las organizaciones deben diseñar un plano de control ligero para la coordinación de los agentes. Este plano de control incluye lo siguiente:

- Registros de capacidades que definen lo que puede hacer cada agente y admiten metadatos versionados para su detección por pares
- Lógica de arbitraje de tareas que utiliza agentes árbitros o supervisores para enrutar las tareas en función del contexto, la disponibilidad y la política
- Seguimiento del ciclo de vida y del estado que permite contextualizar las decisiones en tiempo real y realizar transferencias seguras

Los planos de control garantizan que los sistemas multiagente sigan siendo ampliables, estén alineados con las políticas y sean tolerantes a los fallos, sin centralizar la autoridad ni ralentizar las operaciones.

Sin embargo, los entornos con varios agentes también plantean desafíos operativos. Mantener el contexto en todas las interacciones entre los agentes, gestionar el estado compartido y coordinar las acciones pueden aumentar la complejidad y los costes. Los costes pueden aumentar si utilizas LLMs esos tokens de consumo durante la comunicación entre agentes. Estos costes deben sopesarse frente a los beneficios empresariales combinados que supone la autonomía inteligente a gran escala.

Para abordar estos desafíos, considere la posibilidad de utilizar plataformas de agencias que resuman las principales preocupaciones, como las siguientes:

- Protocolos de comunicación y formatos semánticos estandarizados
- Lógica de orquestación integrada y enrutamiento dinámico
- Administración de memoria y contexto compartidos entre agentes
- Gestión de las alternativas y degradación ordenada durante los fallos

Para los equipos que adoptan estrategias con múltiples agentes, el mejor enfoque es empezar poco a poco y diseñar a medida. Comience con soluciones específicas para un solo agente que resuelvan problemas reales. Luego, componga progresivamente estos agentes en un sistema cooperativo en el que cada uno pueda descubrir, coordinar y delegar en función de los objetivos compartidos y del contexto de todo el sistema.

Es importante destacar que el manejo robusto de los errores y la degradación ordenada deben ser los principios fundamentales del diseño. Los sistemas multiagente deben poder continuar con flujos de trabajo parciales o iniciar una lógica de respaldo cuando los agentes no están disponibles o fallan. Esto promueve la confiabilidad sin un acoplamiento rígido.

Servicios de AWS ofrecen funciones sólidas para respaldar esta arquitectura a escala. [Amazon EventBridge](#) y [EventBridge Pipes](#) proporcionan la columna vertebral estructurada y basada en eventos para la mensajería multiagente. Para gestionar el comportamiento modular, [AWS AppConfig](#) permite una configuración dinámica y segura que permite alternar entre las instancias de los agentes. Para respaldar la administración de memoria y contexto compartidos, utilice [Amazon DynamoDB](#) para lograr una persistencia de estados ligera y adaptada a los inquilinos y una rápida recuperación del contexto entre los agentes. Puede utilizar [Amazon Simple Storage Service \(Amazon S3\)](#) para almacenar historiales de solicitudes estructurados, artefactos compartidos o salidas generadas por agentes. Para flujos de trabajo más complejos que requieren una coordinación

constante, [AWS Step Functions](#) puede organizar procesos de larga duración con puntos de control y lógica de recuperación de errores. En conjunto, estos servicios le ayudan a crear sistemas multiagente componibles, resilientes y conectados semánticamente que se adaptan a las demandas empresariales.

El valor empresarial de los sistemas multiagente

Si bien muchas organizaciones comienzan su viaje hacia la IA con soluciones de un solo agente, todo el potencial de la IA de agente se desbloquea mediante sistemas multiagente escalables. Estos sistemas son fundamentales para resolver problemas complejos y distribuidos y crear ecosistemas de IA sólidos y flexibles que evolucionen con las necesidades empresariales.

Entre las principales ventajas empresariales de los sistemas multiagente se incluyen las siguientes:

- **Escalabilidad:** las tareas y las cargas de trabajo se pueden distribuir entre agentes especializados para aumentar la capacidad y el rendimiento.
- **Flexibilidad:** los agentes se pueden añadir, reemplazar o modificar con una interrupción mínima, lo que permite la agilidad en entornos dinámicos.
- **Resiliencia:** la estabilidad del sistema se mantiene incluso cuando los agentes individuales fallan, gracias a las funciones redundantes y a la conmutación por error inteligente.
- **Especialización:** los agentes diseñados específicamente realizan tareas con mayor eficiencia y precisión.
- **Rentabilidad:** los componentes de los agentes reutilizables aceleran el desarrollo y reducen el coste del despliegue de nuevas capacidades.

Si bien los sistemas multiagente requieren una planificación más temprana, ofrecen agilidad, velocidad y capacidad de innovación a largo plazo. Las empresas que invierten en arquitecturas flexibles de colaboración entre agentes están en condiciones de implementar nuevas capacidades de IA con rapidez, adaptarse a las demandas cambiantes y liderar en un panorama competitivo cada vez más impulsado por los agentes.

Área de enfoque 3: Diseñar soluciones para el control y la multitenencia

Trabajo pendiente: «Ayúdeme a escalar el uso de agentes entre varios clientes sin perder el control, la responsabilidad o la visibilidad».

Los primeros prototipos sirven para demostrar su valor de forma aislada, pero la mayoría de las empresas necesitan dar soporte simultáneamente a varios clientes, departamentos o flujos de trabajo. Esto significa que cada agente debe operar dentro de unos límites de política, datos e identidad claramente definidos. Sin la multitenencia, las operaciones se vuelven frágiles y costosas, y la gobernanza se convierte en un mosaico.

Strategy (Estrategia)

Siga los principios de las arquitecturas de software como servicio (SaaS). Por ejemplo, diseñe para aislar a los inquilinos, aplicar políticas y controlar los recursos. Diseñe agentes y plataformas de orquestación con memoria, configuración e identidad adaptadas a los inquilinos. Para hacer cumplir los límites, utilice el etiquetado, el control de acceso basado en roles (RBAC) y el alcance de la administración de identidades y accesos.

Adopte una capa de observabilidad unificada en la que la telemetría de los agentes se agregue según el contexto del inquilino. Implemente motores de políticas centralizados y una alternancia de capacidades basada en la configuración para hacer cumplir las reglas de comportamiento dinámico.

Cree el despliegue de agentes como un servicio. Permita que los equipos internos o los clientes utilicen las capacidades de los agentes de forma escalable y controlada APIs. AWS proporciona una base sólida para estos patrones. Puede usar [Amazon Cognito para](#) administrar la identidad de los usuarios y los inquilinos, [AWS Organizations](#) y [las políticas de control de servicios \(SCPs\)](#) para el gobierno entre cuentas y [AWS Resource Access Manager \(AWS RAM\)](#) para compartir capacidades de forma segura. Además, [AWS AppConfig](#) puede gestionar de forma dinámica el comportamiento de los agentes por inquilino o entorno. Estos servicios ayudan a hacer cumplir los límites y las políticas, al tiempo que respaldan la infraestructura compartida.

Esta transición del despliegue estático al aprovisionamiento dinámico convierte a la IA de los agentes en una plataforma para toda la empresa.

El valor empresarial de las plataformas de agentes multiusuario

La multitenencia es más que una comodidad arquitectónica: es un acelerador de negocios. A medida que los agentes inteligentes proliferan en los departamentos y equipos, las organizaciones deben respaldar el crecimiento sin duplicar la infraestructura ni fragmentar la gobernanza.

Entre las principales ventajas empresariales de los sistemas multiusuario se incluyen las siguientes:

- **Escalabilidad:** una plataforma de agentes multiusuario permite a los equipos internos, las unidades de negocio o los clientes incorporar las capacidades de IA con mayor rapidez sin necesidad de entornos personalizados.
- **Rentabilidad:** la infraestructura compartida minimiza las implementaciones redundantes, consolida los costos operativos y simplifica el mantenimiento en todos los entornos.
- **Gobernanza y reducción de riesgos:** los controles de políticas centralizados, los modelos de identidad y la observabilidad ayudan a los agentes a operar de forma más segura y de conformidad con las normas en todos los inquilinos.
- **Reutilización de los servicios:** para promover la reutilización y reducir la duplicación, los agentes que prestan atención a los inquilinos pueden ofrecerse como servicios internos, por ejemplo, con fines de enriquecimiento, cumplimiento o resumen.

Entre los ejemplos de casos de uso de sistemas multiusuario se incluyen los siguientes:

- Un agente de cumplimiento que se despliega en todas las filiales adapta su lógica a las normativas locales mediante una configuración específica para cada inquilino. Esto elimina la necesidad de crear agentes independientes para cada región.
- Un agente de automatización del flujo de trabajo interno atiende a varios departamentos con diferentes límites de datos y permisos. Mantiene el aislamiento y, al mismo tiempo, acelera el cumplimiento de las tareas.

Al diseñar los agentes como multi-tenant-aware servicios, las organizaciones evitan la sobrecarga de las iniciativas de IA aisladas. En cambio, fomentan una plataforma de inteligencia unificada. Esta arquitectura permite una implementación escalable, coherencia operativa y un mejor ROI. También facilita la expansión de la adopción de la IA en toda la empresa.

Área de enfoque 4: Generar confianza a través de la identidad, las barreras y la observabilidad

Trabajo por hacer: «Deme la confianza de que los agentes actuarán de manera segura y predecible, especialmente cuando nadie esté mirando».

Los agentes autónomos desafían los modelos de control tradicionales. Su capacidad para razonar y actuar de forma independiente supone un riesgo si no se gestionan adecuadamente. Sin restricciones claras de propiedad, auditabilidad o políticas, pueden desviarse del comportamiento

previsto. Generar confianza en la organización requiere algo más que confiabilidad técnica. Exige explicabilidad, responsabilidad y coherencia.

Strategy (Estrategia)

Cree un sistema de control que priorice la identidad como columna vertebral de una autonomía confiable. Cada agente debe operar con una identidad verificable, permisos específicos y un historial de ejecución rastreable. Los agentes deben estar integrados en un [marco de confianza cero](#) que incluya la vinculación de inquilinos, la herencia del acceso contextual y la aplicación del tiempo de ejecución mediante barreras y motores de políticas. Esto le permite auditar, revertir o restringir las acciones de los agentes en función de las reglas organizativas y la postura de riesgo.

Integre la aplicación de la confianza en tiempo de ejecución mediante barreras inteligentes. Esto incluye controles de velocidad y regulación basados en patrones de comportamiento o condiciones de carga de trabajo, límites de recursos impuestos junto con el autoscalamiento y puntuación de decisiones para evaluar el riesgo. Cree activadores para activar los human-in-the-loop flujos de trabajo cuando se superen los umbrales.

Todos los agentes también deben ser transparentes y explicables. Incorpore una telemetría estructurada mediante registros, rastreos y resúmenes de razonamiento para exponer la lógica de las decisiones. Support Decision Trails y el rastreo de impactos. Esto le ayuda a vincular las acciones de los agentes con las métricas o los resultados clave. Implemente mecanismos de detección de desviaciones que supervisen las desviaciones con respecto al comportamiento o las políticas esperados.

Introduzca agentes reflectantes que observen continuamente el comportamiento de los agentes y los patrones del sistema. Deberían detectar anomalías o inconsistencias en tiempo real. Estos agentes contribuyen a los circuitos de retroalimentación de la gobernanza que pueden iniciar la revalidación, la adaptación o el desmantelamiento de las capacidades.

Establezca juntas de gobierno que revisen las políticas de los agentes, aprueben los cambios de capacidad y supervisen los protocolos de respuesta a los incidentes. La confianza debe ganarse, medirse y reforzarse continuamente.

AWS proporciona una base sólida para implementar este marco de confianza:

- [AWS Identity and Access Management \(IAM\) impone límites](#) de permisos y ejecución basados en roles
- [Amazon CloudWatch](#) y [AWS X-Ray](#) Support tienen total visibilidad y trazabilidad.

- [Amazon GuardDuty](#) y [AWS Config](#) detecte anomalías de seguridad o desviaciones políticas.

En conjunto, estos servicios permiten la aplicación de la identidad, la seguridad en el tiempo de ejecución y la gobernanza basada en la confianza a escala. Pueden ayudar a que los sistemas autónomos sean potentes y confiables.

El valor empresarial de una autonomía fiable

A medida que los agentes se vuelven más autónomos, la confianza se convierte en un factor fundamental para la adopción, la gobernanza y el rendimiento operativo de las empresas. Establecer una base de identidad, observabilidad y barreras ayuda a las organizaciones a escalar la IA de los agentes a dominios sensibles, sin sacrificar la gobernanza ni el control.

Entre los principales impulsores empresariales se incluyen los siguientes:

- **Garantía de gobernanza:** los modelos de identidad sólidos, los registros de auditoría y los límites de los permisos reducen el riesgo de cumplimiento y respaldan la alineación normativa.
- **Continuidad operativa:** las barreras de seguridad en tiempo de ejecución y la detección de anomalías ayudan a prevenir comportamientos no deseados y facilitan la autorrecuperación en casos extremos.
- **Confianza de las partes interesadas:** la explicabilidad de las decisiones y la telemetría generan confianza entre las partes interesadas internas, los gestores de riesgos y los auditores externos.
- **Resiliencia ante los incidentes:** la observabilidad integrada acelera el análisis de la causa raíz y el tiempo de respuesta cuando surgen problemas.

Entre los casos de uso de ejemplo se incluyen:

- En los servicios financieros, los agentes de detección de fraudes deben exponer sus razonamientos, registrar cada acción con una identidad rastreable y operar bajo funciones de IAM muy específicas.
- En el sector de la salud, los agentes de clasificación autónomos deben hacer cumplir los controles de seguridad durante el tiempo de ejecución, someterlos a una revisión humana cuando se alcancen los umbrales y proporcionar registros completos para la supervisión clínica.

Al integrar los mecanismos de confianza en el ciclo de vida de los agentes, las organizaciones pueden permitir que sus sistemas funcionen de forma autónoma y responsable. Esta base reduce el riesgo y permite a los agentes actuar en nombre de la empresa con transparencia e integridad.

En última instancia, una autonomía fiable acelera la adopción al ofrecer a los usuarios y a los líderes la confianza necesaria para ampliar los agentes inteligentes a todas las operaciones principales.

Área de enfoque 5: Gestionar el ciclo de vida

Trabajo por hacer: «Asegurarme de que mi equipo pueda mejorar a los agentes con el tiempo, sin caos ni heroísmo».

A diferencia de las aplicaciones tradicionales, que solo se modelan mediante código, el comportamiento de los agentes también depende de las indicaciones, la memoria, las herramientas y el contexto de formación. Estos factores varían con el tiempo. La desviación erosiona la confiabilidad, aumenta los costos y hace que la depuración sea casi imposible. Sin controles del ciclo de vida, los agentes dejan de ofrecer valor y comienzan a acumular riesgos.

Strategy (Estrategia)

DevOps Establézcalo para los agentes (AgentOps) como práctica. Integre CI/CD canalizaciones diseñadas para los agentes. Utilice estas canalizaciones para probar los resultados rápidos, validar las integraciones de herramientas y perfilar el comportamiento de costo-rendimiento. Mantenga los historiales de versiones de las solicitudes, las políticas y las interacciones de los modelos.

Utilice los circuitos de retroalimentación a partir de los datos de observabilidad para iniciar el readiestramiento, el ajuste rápido o la retirada de los agentes. Incorpore mecanismos de reflexión en todo el sistema, como un registro de mejoras, para institucionalizar el aprendizaje.

Cree un panel de telemetría de rendimiento que muestre la precisión, la latencia, el costo y la confiabilidad de las decisiones. Para agilizar y acelerar la gestión del ciclo de vida mediante la AWS infraestructura, los equipos pueden utilizar los kits de herramientas para agentes. Un ejemplo es el [SDK de Strands Agents](#), que proporciona herramientas estructuradas para un rápido control de versiones, registro de herramientas e integración de CI/CD con Servicios de AWS, como, y. [AWS CodePipeline](#)[AWS Cloud Development Kit \(AWS CDK\)](#)[AWS Lambda](#) Además, utilice [Amazon S3](#) y [Amazon Elastic File System \(Amazon EFS\)](#) para almacenar artefactos de agentes y datos de entrenamiento. Úselo [AWS Step Functions](#) para automatizar flujos de trabajo complejos de reentrenamiento o validación. Puede usar [Amazon SageMaker AI](#) cuando los agentes necesiten

ajustar modelos personalizados o ajustar los flujos de trabajo más allá de la orquestación LLM. La disciplina del ciclo de vida transforma a los agentes, que pasan de ser experimentos a activos duraderos y en evolución.

Con el tiempo, este sistema de ciclo de vida constituye la columna vertebral de la innovación. Le ayuda a recomponer, reorganizar y reimplementar las capacidades con agilidad. Esto transforma la capa de agentes en un sistema vivo, capaz de evolucionar en respuesta tanto a los comentarios como a las oportunidades.

El valor empresarial de la gestión del ciclo de vida

La gestión eficaz del ciclo de vida es un factor clave del rendimiento y la rentabilidad de los agentes. Garantiza que los agentes inteligentes sigan ofreciendo resultados precisos, fiables y acordes con el valor a medida que evolucionan. Los agentes no siguen siendo valiosos por defecto. Deben evolucionar en sincronía con los cambiantes requisitos empresariales, los flujos de trabajo y los entornos de datos. Un AgentOps equipo disciplinado ayuda a los agentes a mantenerse precisos, eficientes y alineados con los objetivos empresariales a lo largo del tiempo.

Entre los principales factores empresariales se incluyen los siguientes:

- **Consistencia del rendimiento:** las pruebas continuas, la rápida validación y el reentrenamiento ayudan a los agentes a mantener la calidad de las decisiones en condiciones y conjuntos de datos cambiantes.
- **Optimización de costes:** la elaboración de perfiles mediante telemetría identifica las herramientas ineficientes, los avisos de alto nivel o las ejecuciones innecesarias. Luego, puede realizar ajustes para reducir los costos operativos.
- **Iteración más rápida:** la automatización del ciclo de vida CI/CD acelera los ciclos de desarrollo, lo que ayuda a los equipos a experimentar, implementar y mejorar los agentes con confianza.
- **Reducción de riesgos:** el rápido control de versiones, la compatibilidad con la reversión y los mecanismos de evaluación estructurados ayudan a prevenir las regresiones y a permitir una gestión de cambios segura y fiable.

Algunos ejemplos de casos de uso incluyen lo siguiente:

- Un agente de atención al cliente supervisa la latencia, el coste del modelo y los comentarios de los usuarios. La observabilidad revela un aumento de los costes, lo que hace que se reajusten las indicaciones integradas y la lógica del modelo alternativo.

- Un agente de resumen de contratos se actualiza en función de los comentarios de los equipos legales. Las instrucciones versionadas se prueban en entornos aislados antes de su lanzamiento a la fase de producción, lo que garantiza su seguridad y calidad.

Con una gestión estructurada del ciclo de vida, las organizaciones van más allá del mantenimiento reactivo y pasan a la mejora proactiva y continua. Los agentes se convierten en activos digitales adaptables que se miden, refinan y revalidan en función de los objetivos empresariales. Esta práctica transforma los ecosistemas de agentes en sistemas de alto rendimiento, resilientes y respetuosos con los costes que ofrecen un valor duradero y, al mismo tiempo, se mantienen al día con los cambios.

Área de enfoque 6: Alinee los modelos de los agentes con los modelos de negocio

Trabajo por hacer: «Muéstreme el impacto para que pueda justificar la inversión continua».

Incluso los agentes técnicamente capaces se convierten en pasivos si no están vinculados a los resultados empresariales. Los agentes deben contribuir a la eficiencia, la monetización o la diferenciación estratégica. Sin embargo, la mayoría de las empresas se esfuerzan por definir cómo se adaptan los agentes a los modelos de precios, paquetes o uso. Sin una alineación clara con el valor empresarial, es difícil justificar la ampliación o incluso el mantenimiento de la inversión.

Strategy (Estrategia)

Adopte prácticas de gestión de productos. Trate a los agentes como servicios monetizables con un ROI mensurable. Defina las estrategias de precios en función de las decisiones, las sesiones o los resultados. Luego, agrupe las capacidades de los agentes en ofertas escalonadas que estén alineadas con los segmentos de clientes o las unidades de negocio internas.

Para promover la sostenibilidad, las organizaciones deben captar tanto el valor directo como los multiplicadores del crecimiento mediante el despliegue de agentes. Considere la posibilidad de utilizar las siguientes métricas de ROI para medir el valor inmediato:

- Costo por decisión: compare los costos de procesamiento de los agentes con los equivalentes humanos.
- Compresión del tiempo: cuantifique el valor de los ciclos acelerados, como las ventas o aprobaciones más rápidas.

- Reducción de errores: mida los ahorros gracias a la mejora de la precisión, la coherencia y el cumplimiento.

Además de estos beneficios inmediatos, los agentes pueden aprovechar las siguientes oportunidades de crecimiento a largo plazo:

- Acumulación de capacidades: combine los servicios de los agentes para crear soluciones verticales específicas para cada dominio.
- Efectos de red: aumente el valor a través de ecosistemas multiagentes en los que la coordinación se suma a la utilidad.
- Extensión del mercado: genere nuevas fuentes de ingresos a través de servicios consumibles externamente y habilitados por agentes.

Cree circuitos de retroalimentación a partir de métricas empresariales (como el ahorro de costes, el aumento de conversiones o time-to-resolution) para impulsar la evolución continua de los agentes. Analice la telemetría de uso y las puntuaciones de satisfacción de los usuarios para afinar su alineación de valores y las prioridades de la hoja de ruta. Al vincular las capacidades de los agentes directamente con los modelos de negocio, las organizaciones se posicionan para captar un valor acumulable y sostenible, no solo resultados técnicos.

Lo siguiente Servicios de AWS respalda esta alineación al proporcionar marcos sólidos de seguimiento y monetización:

- [AWS Cost Explorer](#) y [Amazon CloudWatch](#) ofrecen información sobre los costes por agente y la eficiencia operativa.
- [Amazon API Gateway](#) permite el acceso medido, la limitación de velocidad y los precios escalonados para los puntos de enlace de los agentes.
- [AWS Marketplace](#) proporciona un canal para los agentes de publicación y las soluciones de agencias como productos comerciales.

Estos servicios le ayudan a transformar la funcionalidad de los agentes en ofertas digitales escalables y orientadas al valor que se alinean con las estrategias de crecimiento y monetización empresarial.

Entrega de software en evolución para la IA de los agentes

La entrega de software moderna se basa en una simple suposición: usted controla los sistemas que envía. Usted define los requisitos, escribe la lógica, comprueba los resultados esperados e implementa servicios predecibles. Incluso los DevOps enfoques ágiles siguen basándose en el principio de que cada sprint ofrece algo determinista, verificable y, en gran medida, sujeto a la supervisión humana.

La IA de las agencias pone patas arriba esa base. Los sistemas de los agentes interpretan, razonan y se adaptan en lugar de seguir los guiones. Su comportamiento depende del código que escribas, del contexto en el que operan, de las entradas que reciben, de las herramientas a las que pueden acceder y de los objetivos que se les asignen. En resumen, no siguen órdenes; persiguen resultados.

Esto hace que la entrega se centre menos en el control y más en la alineación. En lugar de dar instrucciones, debe configurar su comportamiento. Esto significa que el ciclo de vida de desarrollo de software (SDLC) tradicional ya no es adecuado porque se diseñó para sistemas controlados por humanos y basados en la lógica.

Esta sección contiene los siguientes temas:

- [Zonas de intención para la IA de los agentes](#)
- [Evolución del ciclo de vida de entrega de la IA de los agentes](#)
- [Preparar a los equipos para la IA de los agentes](#)

Zonas de intención para la IA de los agentes

En lugar de etapas rígidas, como definir, construir, probar y lanzar, necesitamos un modelo que abarque la autonomía, la incertidumbre y la emergencia. En su lugar, utilizas zonas de intención. Una zona de intención define un espacio acotado en el que un agente puede operar con autonomía, dentro de las limitaciones. El objetivo es pasar de la microgestión de todas las tareas a diseñar entornos en los que los agentes puedan actuar, aprender y colaborar de forma segura. Usted especifica el qué (el resultado deseado), el por qué (la intención) y las barreras (las restricciones, las políticas y los límites de confianza). Teniendo en cuenta esos límites y esta información, el agente averigua cómo hacerlo.

En lugar de una línea de montaje, piense en el medio ambiente como un espacio aéreo. Usted controla quién puede entrar, qué puede hacer y adónde puede ir. Pero una vez dentro, pueden navegar libremente según sea necesario. Así es como los sistemas de agentes se escalan sin caos.

No se trata solo de un cambio filosófico, sino práctico. El resultado no determinista de los sistemas basados en agentes no se puede probar completamente mediante pruebas unitarias. No se puede versionar como los binarios estáticos. Los agentes cambian con el tiempo, se adaptan a los nuevos datos e interactúan con otros sistemas de formas impredecibles. Intentar entregarlos mediante modelos tradicionales conduce a arquitecturas frágiles e inescalables. En el peor de los casos, lleva a una falsa confianza en sistemas que en realidad no se pueden gobernar.

Cuando los equipos adoptan la entrega basada en la intención, obtienen dos ventajas:

- Controla donde más importa: definen los límites en lugar de los resultados.
- Escalabilidad mediante la delegación: permiten a los agentes gestionar una complejidad que los humanos no pueden codificar de forma rígida.

Así es como se pasa de prototipos aislados a sistemas de agentes reales, aptos para producción, capaces de ofrecer valor de forma repetida y fiable.

Evolución del ciclo de vida de entrega de la IA de los agentes

Para apoyar un comportamiento inteligente y adaptativo, el SDLC debe replantearse y dejar de ser un control determinista para convertirse en una intención adaptativa. Los siguientes son los cambios necesarios para desarrollar el SDLC tradicional para la IA de los agentes:

- La planificación se convierte en diseño intencional. Los equipos definen los objetivos, las limitaciones y los comportamientos esperados de los agentes. Las políticas y los criterios de éxito se enmarcan en términos de alineación, no de lógica.
- La arquitectura se convierte en andamiaje. Los equipos se centran en definir las funciones, las interfaces, las barreras, los mecanismos alternativos y la observabilidad, en lugar de programar cada ruta de decisión.
- Las pruebas se convierten en una evaluación del comportamiento. En lugar de afirmar resultados específicos, los equipos validan si los agentes se mantienen dentro de los límites aceptables y cumplen su intención con diversas variables.
- El despliegue se convierte en una orquestación continua. Los sistemas de Agentic se implementan con controles de tiempo de ejecución, monitoreo en vivo y canales de retroalimentación que permiten la optimización en tiempo real.
- La iteración se convierte en retroalimentación y adaptación. En lugar de los ciclos tradicionales de parches de cambio de código, los equipos observan cómo evolucionan los agentes, dónde tienen

éxito o cuándo se desvían. Cuando es necesario, los equipos intervienen mediante la actualización de las restricciones, el readiestramiento y la adición o modificación de mecanismos de control.

Las prácticas actuales que se centran en la iteración, la experimentación y la retroalimentación rápida están a mitad de camino. El cambio a los sistemas de agencia no supone un rechazo de los principios ágiles. De hecho, es una evolución natural de los mismos. El pensamiento ágil hace hincapié en la adaptabilidad, la retroalimentación y las soluciones de trabajo por encima de los planes rígidos. Esto se alinea perfectamente con la naturaleza de los sistemas de agentes, que aprenden, se adaptan y responden al contexto en tiempo real. Si ya está ejecutando ciclos cortos, validando las suposiciones rápidamente y gestionando la incertidumbre mediante una entrega continua, está bien preparado para liderar esta transición.

Sin embargo, existen diferencias clave. El enfoque ágil tradicional supone que lo que se entrega es determinista. Supone que, una vez construida, la cosa se comportará de forma coherente y predecible, con resultados repetibles para las mismas entradas. Esta repetibilidad le ayuda a depurar, probar e iterar con confianza. Los sistemas de agencia rompen ese modelo. Son probabilísticos, sensibles al contexto y capaces de evolucionar de forma independiente. Esto significa que algunas prácticas ágiles se vuelven menos útiles, como el seguimiento de la velocidad en función de la finalización de la historia, los estrictos criterios de aceptación o la planificación determinista de los sprints.

Los siguientes aspectos del SDLC tradicional se aplican a la IA de los agentes:

- Desarrollo y entrega iterativos
- Los comentarios de los clientes como señal principal
- Colaboración interfuncional
- Integración e implementación continuas

Los siguientes aspectos del SDLC tradicional deben evolucionar para la IA de los agentes:

- Redefina lo hecho como alineado con la intención. Concéntrese en si el comportamiento del agente satisface el objetivo previsto dentro de las restricciones definidas.
- Cambie de los criterios de aceptación a las barreras conductuales.
- Amplíe la definición de hecho para incluir la preparación para el tiempo de ejecución, lo que incluye mecanismos de observabilidad, explicabilidad y retroalimentación que respaldan el aprendizaje continuo y la confianza.

- Priorice los circuitos de retroalimentación en tiempo real y el seguimiento del comportamiento por encima de la planificación inicial

La buena noticia es que no necesitas tirar a la basura el manual del SDLC. Solo hay que hacer que pase de gestionar el código a dar forma a la conducta. En los sistemas de agencias, el éxito no se basa solo en si el software se ejecuta, sino en cómo se comporta.

Preparar a los equipos para la IA de los agentes

La ingeniería de software no va a desaparecer. Está evolucionando. El trabajo pasa de escribir funciones a dar forma a marcos y mecanismos de control para un comportamiento inteligente. En el mundo de la IA agencial, construir ya no es la parte difícil, sino gestionar el surgimiento. Para la mayoría de los equipos de ingeniería, la evolución parece un cambio de mentalidad más que un salto técnico. En lugar de preguntarse «¿Qué hará el sistema?» la pregunta es: «¿Qué es lo que le hemos permitido perseguir y cómo sabremos si sigue su curso?»

Para los equipos de ingeniería, la evolución hacia la IA como agente requiere los siguientes cambios:

- Un cambio cultural: los equipos deben sentirse cómodos con la incertidumbre y la autonomía en los sistemas que no controlan por completo.
- Nuevas funciones: los diseñadores de intenciones, los evaluadores del comportamiento y los ingenieros de observabilidad se convierten en elementos fundamentales para su desempeño.
- Lenguaje compartido: los equipos necesitan una comprensión clara y compartida de los objetivos, las barreras y las señales de éxito, del mismo modo que antes necesitaban las especificaciones y los casos de prueba.

A medida que vaya madurando la IA generativa, veremos más sistemas de agencias que interactúan con los clientes, los productos y las operaciones. Las organizaciones que tengan éxito no serán las que cuenten con los mejores modelos. Serán las que puedan integrar a los agentes en los flujos de trabajo del mundo real con confianza, control y rapidez. Eso significa que los modelos de entrega y los equipos de ingeniería deben evolucionar juntos. Las zonas de intención te dan la abstracción necesaria para hacerlo. Le ayudan a hacer operativa la autonomía sin renunciar a la responsabilidad. También ofrecen un marco compartido entre los equipos para ayudar a gobernar sistemas que no se pueden codificar de forma rígida.

Para obtener más información sobre cómo preparar a los equipos para la IA de los agentes, consulte la sección [Preparar la empresa para la IA de los agentes a gran escala](#) de esta guía.

Preparar la empresa para la IA agencial a gran escala

A medida que convergen [las áreas de enfoque](#) descritas en esta guía, la IA de los agentes pasa de ser funciones aisladas a convertirse en una capa de inteligencia unificada que puede entenderse como una plataforma de capacidades. Esta plataforma no solo ejecuta tareas. Evoluciona, se adapta y coordina en todos los dominios. Los agentes se convierten en servicios modulares, reutilizables y detectables que aceleran la innovación, reducen la carga cognitiva e impulsan resultados mensurables en toda la empresa. Esta vista de plataforma sienta las bases para una inteligencia escalable integrada en todo el modelo operativo.

La puesta en marcha de la IA de los agentes requiere algo más que el despliegue de agentes inteligentes. Exige una transformación fundamental en la forma en que la empresa organiza los equipos, diseña los procesos y gobierna la tecnología. Al igual que el cambio a la nube o la DevOps redefinición de los modelos operativos, la IA agentic presenta una nueva era de automatización de decisiones, aprendizaje continuo y coordinación autónoma. El éxito depende de alinear los sistemas, las personas y los procesos en torno a esta nueva filosofía operativa.

Esta sección contiene los siguientes temas:

- [Alinear los equipos y los modelos de propiedad](#)
- [Gestionar el cambio y preparar a la organización](#)
- [Diseñando una arquitectura para la interoperabilidad y la colaboración](#)
- [Incorporar la gobernanza a una estructura de agentes](#)
- [Adoptar una mentalidad operativa que dé prioridad a la toma de decisiones](#)
- [Escalar con un propósito y una intención](#)

Alinear los equipos y los modelos de propiedad

El primer paso hacia la madurez es la alineación interfuncional. Las empresas deben establecer AgentOps equipos que incluyan AI/ML profesionales y especialistas en el campo, como arquitectos de sistemas distribuidos, ingenieros de software, propietarios de productos, líderes de cumplimiento y arquitectos de plataformas. Estos equipos son responsables conjuntamente de todo el ciclo de vida de un agente, desde el diseño y la implementación hasta el readiestramiento y la supervisión.

El aprovisionamiento y el lanzamiento de los agentes deben seguir las prácticas nativas de la nube, como el uso de la infraestructura [AWS Cloud Development Kit \(AWS CDK\)](#) y [AWS](#)

[CodePipeline](#) para la infraestructura como código y el despliegue automatizado. Esta estructura fomenta la responsabilidad compartida y acelera la iteración. Del mismo modo que DevOps unifica el desarrollo y las operaciones, AgentOps conecta la inteligencia con la gobernanza y la ejecución.

Para ser eficaces, estos equipos también necesitan un lenguaje compartido. Las partes interesadas de la empresa deben entender [qué son los agentes](#), [cómo operan](#) y [qué resultados generan](#). La formación y la habilitación interna son esenciales. Al desmitificar a los agentes e integrar este modelo mental en las conversaciones cotidianas, las organizaciones permiten una participación más amplia y una innovación más alineada.

Para acelerar el desarrollo y la integración de los agentes que utilizan Servicios de AWS, los equipos pueden adoptar marcos como el [SDK de Strands Agents](#), que ofrece herramientas basadas en CLI para andamiar, configurar y empaquetar agentes. Strands Agents está diseñado para funcionar sin problemas con AWS infraestructuras, como [Amazon Bedrock AWS Lambda](#), [Amazon EventBridge](#), [AWS CDK](#), [the](#) y [AWS CodePipeline](#). Permite la creación y el despliegue rápidos de prototipos y, al mismo tiempo, mantiene los estándares de producción.

Sin embargo, la estructura y las herramientas por sí solas no son suficientes. Ampliar la IA de los organismos requiere una preparación cultural, educativa y de liderazgo deliberada para garantizar que la adopción se arraigue en toda la organización.

Gestionar el cambio y preparar a la organización

Para escalar con éxito la IA de las agencias se necesita algo más que implementar infraestructura o agentes inteligentes. Exige un enfoque estructurado del cambio organizacional. Esto incluye la preparación cultural, el desarrollo de habilidades, los circuitos de retroalimentación basados en métricas y la alineación de los ejecutivos para garantizar que la adopción sea intencional y sostenible.

Fomentar la evolución cultural

- Posicione a los agentes como compañeros de equipo, no como sustitutos, para reducir la resistencia y generar confianza.
- Comuníquese de manera transparente sobre las capacidades y limitaciones de los agentes para establecer expectativas realistas.
- Establezca protocolos de traspaso claros para determinar cuándo los agentes deben transferir las decisiones a una autoridad superior o delegar partes del proceso en un colaborador humano.

Establezca un marco de desarrollo de habilidades

- Ofrezca formación basada en funciones y adaptada a ingenieros, gerentes de producto, líderes de dominio y funcionarios de cumplimiento.
- Cree centros de excelencia para compartir las mejores prácticas, los patrones de herramientas y los activos reutilizables.
- Une a especialistas en IA con expertos en el campo a través de programas de tutoría para cerrar las brechas de conocimiento.

Defina las métricas y los ciclos de retroalimentación

- Fije el valor técnico y empresarial KPIs al estratégico para evaluar el impacto. Algunos ejemplos de valor son la latencia de decisión, la precisión de la resolución y el ahorro de costes.
- Capture de forma sistemática y continua los comentarios de los usuarios sobre los puntos de fricción y los desafíos de adopción.
- Realice retrospectivas periódicas para evaluar el rendimiento de los agentes, las tendencias de uso y las oportunidades de mejora.

Alinee el liderazgo desde arriba

- Obtenga el patrocinio ejecutivo vinculando las iniciativas de los agentes con los resultados estratégicos y el ROI.
- Forme comités de gobierno multifuncionales que incluyan líderes técnicos y empresariales.
- Adapte las estrategias de comunicación para lograr claridad y compromiso en todos los niveles de la organización.

Este enfoque sistemático de la gestión del cambio garantiza que la implementación de la tecnología vaya acompañada de la madurez organizacional. Crea una base para la confianza, la adopción y el valor empresarial a largo plazo.

Diseñando una arquitectura para la interoperabilidad y la colaboración

Los despliegues de agentes aislados ofrecen beneficios a nivel local. Sin embargo, el valor empresarial surge cuando los agentes pueden descubrir, invocar y colaborar entre sí de forma

dinámica. Esto implica definir estándares para el registro, la autenticación y el intercambio de capacidades de los agentes. Desde el punto de vista arquitectónico, esto refleja el cambio de los monolitos a los microservicios, que son unidades componibles, reutilizables y ligeramente acopladas que resuelven problemas complejos de forma conjunta.

[Los protocolos emergentes, como el A2A y el MCP, son fundamentales.](#) Permiten la interoperabilidad semántica entre agentes, herramientas y sistemas de memoria. El A2A admite la interacción entre pares, lo que permite a los agentes negociar la propiedad de las tareas, compartir el contexto y coordinar los flujos de trabajo. MCP complementa esto al ofrecer esquemas compartidos para intercambiar datos contextuales entre los agentes y sus entornos. Estandariza la forma en que se invocan las funciones, se accede a ellas y APIs se mantienen los estados. En conjunto, estos protocolos promueven la extensibilidad, la coherencia y el mantenimiento a largo plazo en todo el ecosistema de agentes.

La gobernanza sigue siendo fundamental. Las capas de control, como los agentes árbitros, permiten delegar en función de las políticas sin introducir cuellos de botella centralizados. Estos agentes actúan como intermediarios de confianza. Hacen cumplir los límites y permiten que otros agentes se autoorganicen. La colaboración entre agentes ayuda a las organizaciones a escalar sus ecosistemas de inteligencia artificial con agilidad y confianza.

Incorporar la gobernanza a una estructura de agentes

Una mayor autonomía conlleva un mayor riesgo. La gobernanza debe estar integrada en la arquitectura de los agentes desde el primer día. Esto incluye definir los límites de las políticas que determinen lo que los agentes pueden hacer, aplicar modelos de identidad que determinen en nombre de quién actúan e implementar la explicabilidad y la trazabilidad. Los sistemas de observabilidad deben capturar la telemetría del comportamiento de los agentes mediante servicios como [Amazon CloudWatch](#) y [AWS X-Ray](#), que proporcionan un registro centralizado y un seguimiento distribuido en todos los flujos de trabajo de los agentes. Los agentes reflexivos pueden auditar y evaluar el rendimiento de forma continua en función de estas fuentes de telemetría.

La gobernanza también debe evolucionar a medida que el ecosistema de agentes vaya madurando. A medida que los agentes se vuelven más capaces y autónomos, los mecanismos de supervisión deben adaptarse mejor. Las actualizaciones de las políticas, la limitación de capacidades y las restricciones de comportamiento durante el tiempo de ejecución deben ser dinámicas y aplicables a escala. La confianza no es una característica adicional. Se refuerza continuamente a través de la arquitectura, el comportamiento y el proceso. [AWS Identity and Access Management \(IAM\)](#) y [AWS AppConfig](#) desempeñan un papel fundamental a la hora de garantizar la seguridad de las identidades,

los límites de los permisos de ejecución y los cambios de comportamiento específicos del entorno entre los agentes.

Adoptar una mentalidad operativa que dé prioridad a la toma de decisiones

La automatización tradicional se centra en la eficiencia de los procesos, que consiste en ejecutar scripts o flujos de trabajo predefinidos de forma más rápida y fiable. La IA de Agentic, por el contrario, introduce la automatización que prioriza la toma de decisiones. Los agentes evalúan el contexto, sopesan las opciones y adaptan el comportamiento en tiempo real. Este cambio de una mentalidad que prioriza la ejecución a otra que es lo primero en la toma de decisiones requiere una nueva forma de pensar en las métricas de éxito y los resultados. En lugar de medir el éxito exclusivamente en función de la finalización de las tareas, el éxito de la IA de los agentes se mide en función del grado de coherencia de la decisión con la intención, las políticas y las condiciones cambiantes.

En lugar de medir únicamente la finalización de las tareas o el tiempo del ciclo, las organizaciones deben evaluar la calidad de las decisiones y la capacidad de respuesta al cambio. time-to-action KPIs debe incluir métricas como:

- Calidad de la decisión: ¿en qué medida personalizó el agente su respuesta al usuario o escenario específico? ¿Tomó decisiones matizadas que estaban alineadas con los objetivos empresariales y el contexto del usuario?
- Time-to-action — ¿Con qué rapidez e inteligencia evaluó el agente una situación y respondió? ¿Era la latencia lo suficientemente baja como para dar una sensación adaptativa y parecida a la humana?
- Descarga cognitiva: ¿cuánto análisis manual, clasificación o toma de decisiones rutinarias pudo realizar el agente en nombre de un humano? ¿Redució el esfuerzo o simplemente lo modificó?

Las empresas que adoptan la mentalidad de dar prioridad a la toma de decisiones pueden ser más resilientes, adaptables y capaces de operar en un nuevo nivel de complejidad.

Escalar con un propósito y una intención

Escalar con éxito la IA de los agentes no consiste en experimentar con más herramientas. Se trata de crear una capa duradera de inteligencia empresarial. Esto requiere inversiones en la

infraestructura de la plataforma, la cultura operativa, los marcos de gobierno y la alineación estratégica. Las empresas deben adoptar un enfoque intencional. Deben tratar a los agentes no como experimentos sino como componentes fundamentales de su modelo operativo digital.

La alineación con el marco de [AWS Well-Architected](#) ayuda a sus sistemas a cumplir con los estándares empresariales de confiabilidad, seguridad, eficiencia del rendimiento y optimización de costos. Herramientas como el [SDK de Strands Agents](#) pueden acelerar este proceso al proporcionar instrucciones estructuradas, registro de herramientas y preparación para la CI/CD. Esto ayuda a los equipos a pasar de la experimentación a la entrega escalable mediante el uso de flujos de trabajo conocidos. AWS

La IA de los agentes no es una herramienta, sino un cambio en la forma en que la inteligencia se integra en las operaciones. Las organizaciones que se preparan en consecuencia pueden automatizar más, operar de manera más inteligente, adaptarse más rápido y crear una ventaja duradera en un mundo cada vez más complejo.

Conclusión sobre la operacionalización de la IA de los agentes

La IA de las agencias representa más que un cambio tecnológico. Marca el surgimiento de un nuevo sistema operativo para la empresa. Las organizaciones que adoptan esta transformación van más allá de los casos de uso limitados de la automatización y convierten la inteligencia en la base de sus operaciones. Este cambio consiste en rediseñar la forma en que se toman las decisiones, cómo se adaptan los sistemas y cómo se obtienen los resultados a escala.

En una era definida por la creciente complejidad, la demanda en tiempo real y la sobrecarga de información, el modelo tradicional de automatización mediante scripts ha llegado a sus límites. El éxito ahora depende de la capacidad de integrar la inteligencia directamente en los flujos de trabajo para crear sistemas que perciban, razonen, actúen y evolucionen. La IA de las agencias puede alinear la autonomía con el propósito, la toma de decisiones con la gobernanza y la adaptabilidad con la responsabilidad.

Esta transición requiere pasar de pensar primero en la ejecución a pensar primero en la toma de decisiones. Los sistemas de agencia no se limitan a seguir instrucciones. Interpretan los objetivos, sopesan las compensaciones y persiguen los resultados dentro de unas restricciones definidas. En este contexto, el éxito se mide no solo por la finalización de las tareas. También se mide por la calidad, la agilidad y la explicabilidad de las decisiones que se toman en tiempo real. Las organizaciones deben replantearse las métricas, los incentivos y el diseño del sistema para apoyar a los agentes que operan de forma inteligente en situaciones de incertidumbre.

Operacionalizar la IA de los agentes no es una mejora. *plug-and-play* Es una transformación arquitectónica y cultural. Requiere prácticas disciplinadas en la gestión del ciclo de vida, la aplicación de la confianza, la interoperabilidad y la alineación con los modelos empresariales. También es necesaria la evolución de los modelos de entrega, por ejemplo, la configuración de las zonas de intención, la incorporación de barreras de tiempo de ejecución y la alineación continua del comportamiento de los agentes con los resultados estratégicos. Los equipos deben adoptar un lenguaje común, una propiedad y una responsabilidad compartidas en lo que respecta al rendimiento y la seguridad de los agentes.

La preparación empresarial puede determinar quién prospera en este nuevo entorno. Las organizaciones deben invertir en la habilitación interna, AgentOps las capacidades y los marcos de gobierno que escalen y creen valor a largo plazo. Las personas que tienen éxito pueden

crear sistemas más inteligentes y también pueden crear empresas más adaptables, resilientes e impulsadas por la información.

Esta guía sienta las bases. Conecta la estrategia con la ejecución y prepara a las organizaciones para crear plataformas escalables de agentes inteligentes. La serie de contenido más amplia sobre la IA de los agentes AWS proporciona orientación complementaria. Para ver las demás guías de esta serie, consulte [Agentic AI](#) en el sitio web de orientación prescriptiva. AWS Esta serie de contenido ofrece una hoja de ruta para poner en práctica la autonomía con disciplina e intención.

Para empezar, identifique un espacio de toma de decisiones de alto impacto en el que los agentes puedan ofrecer mejoras mensurables en cuanto a velocidad, precisión o capacidad de respuesta. Luego, despliegue un agente piloto específico que cuente con instrumentación, gobierno y circuitos de retroalimentación. Úselo para validar la hipótesis del valor, generar un impulso interno y generar confianza en el enfoque. El impulso se agrava a través del aprendizaje.

La IA de las agencias no es un destino; es una capa de capacidades que evoluciona junto con su negocio. Representa un cambio a largo plazo hacia la inteligencia como infraestructura. Las organizaciones que lideran en este ámbito pueden automatizar más, responder más rápido, adaptarse mejor y crear modelos operativos que sean capaces de sortear la complejidad a escala empresarial.

Recursos para poner en funcionamiento la IA de los agentes

Servicios de AWS

Las siguientes características Servicios de AWS y características pueden ayudarle a crear y poner en funcionamiento los sistemas de IA de los agentes en: Nube de AWS

- [Amazon API Gateway](#) puede exponer las capacidades de los agentes como escalables y ofrece precios basados en el uso.
- [AWS AppConfig](#) ofrece gestión de la configuración en tiempo de ejecución y alternancia de funciones para los agentes de distintos inquilinos o entornos.
- [Amazon Bedrock](#) es un servicio modelo básico que los agentes pueden utilizar para razonar, generar y ejecutar rápidamente.
- [AWS Cloud Development Kit \(AWS CDK\)](#) es un servicio de infraestructura como código que puede utilizar para implementar y gestionar pilas de agentes.
- [AWS CloudTrail](#) registra el historial de eventos para que pueda realizar un seguimiento de la actividad de los agentes, los registros de auditoría y los comportamientos de integración.
- [Amazon CloudWatch](#) puede gestionar registros, métricas y alarmas para supervisar el rendimiento de los agentes y el comportamiento de colaboración entre varios agentes.
- [AWS CodePipeline](#) proporciona CI/CD automatización que puede utilizar para probar, validar e implementar el código de los agentes.
- [Amazon Cognito](#) es un servicio de identidad que puede utilizar para administrar la autenticación de usuarios y arrendatarios en sistemas con varios agentes.
- [AWS Config](#) ofrece detección de conformidad y de desviaciones para la configuración del entorno y la política de los agentes.
- [AWS Cost Explorer](#) puede realizar un seguimiento del uso a nivel de agente y ayudar a alinear los costes para maximizar el ROI.
- [Amazon DynamoDB](#) es un servicio de almacenamiento que puede utilizar para la memoria de los agentes, los registros de mejoras y el estado contextual.
- [Amazon Elastic File System \(Amazon EFS\)](#) es un sistema de archivos compartidos que puede utilizar para la colaboración entre agentes o para el procesamiento intermedio entre flujos de trabajo.

- [Amazon EventBridge](#) es un bus de eventos principal que puede utilizar para enrutar tareas y organizar la comunicación en la estructura de agentes.
- [Amazon EventBridge Pipes](#) puede agilizar la gestión y el enrutamiento de eventos para conectar agentes y servicios.
- [Amazon GuardDuty](#) ofrece detección de amenazas y monitoreo de anomalías que pueden respaldar la ejecución segura de los agentes.
- [AWS Identity and Access Management \(IAM\)](#) le ayuda a definir permisos detallados para la ejecución de los agentes y el acceso a los datos.
- [AWS Lambda](#) es un servicio informático sin estado que puede ejecutar la lógica de los agentes y enjamburar drones.
- [AWS Marketplace](#) es una plataforma de distribución externa que puede utilizar para ofrecer funciones de agente como productos comerciales.
- [AWS Organizations](#) es un servicio de gobernanza y aplicación de políticas multicuenta que puede ayudarle a gestionar la infraestructura de agentes multiusuario.
- [AWS Organizations las políticas de control de servicios](#) actúan como barreras para controlar los permisos a nivel de cuenta o unidad organizativa.
- [Amazon Quick](#) es una plataforma generativa de inteligencia empresarial (BI) basada en IA que le ayuda a analizar datos, crear visualizaciones, automatizar flujos de trabajo y colaborar con otras personas de su organización.
- [AWS Resource Access Manager \(AWS RAM\)](#) puede ayudarlo a compartir capacidades entre las cuentas y los servicios de los agentes.
- [Amazon SageMaker AI](#) es un servicio que puede utilizar para el entrenamiento, el ajuste y la inferencia de modelos más allá de los modelos fundamentales.
- [Amazon Simple Storage Service \(Amazon S3\)](#) ofrece almacenamiento de objetos para bibliotecas de solicitudes, artefactos de modelos y datos generados por agentes.
- [AWS Step Functions](#) es un motor de flujo de trabajo que puede ayudarle a coordinar los flujos de varios agentes y los procesos de reentrenamiento.
- [AWS X-Ray](#) ofrece un rastreo distribuido que puede usar para rastrear los flujos de decisiones de los agentes y las dependencias de los servicios.

Otros recursos AWS

- [Fundamentos de la IA agencial en AWS](#)

- [Los patrones y flujos de trabajo de la IA de los agentes están activos AWS](#)
- [Los marcos, protocolos y herramientas de inteligencia artificial de las agencias están disponibles AWS](#)
- [Creación de arquitecturas sin servidor para la IA de los agentes en AWS](#)
- [Creación de arquitecturas multiusuario para la IA de los agentes en AWS](#)

Historial de documentos

En la siguiente tabla, se describen cambios significativos de esta guía. Si quiere recibir notificaciones de futuras actualizaciones, puede suscribirse a las [notificaciones RSS](#).

Cambio	Descripción	Fecha
Publicación inicial	—	12 de agosto de 2025

AWS Glosario de orientación prescriptiva

Los siguientes son términos de uso común en las estrategias, guías y patrones proporcionados por la Guía AWS prescriptiva. Para sugerir entradas, utilice el enlace [Enviar comentarios](#) al final del glosario.

Números

Las 7 R

Siete estrategias de migración comunes para trasladar aplicaciones a la nube. Estas estrategias se basan en las 5 R que Gartner identificó en 2011 y consisten en lo siguiente:

- **Refactor/re-architect** — Mueva una aplicación y modifique su arquitectura aprovechando al máximo las funciones nativas de la nube para mejorar la agilidad, el rendimiento y la escalabilidad. Por lo general, esto implica trasladar el sistema operativo y la base de datos. Ejemplo: migre su base de datos Oracle local a la PostgreSQL-Compatible edición Amazon Aurora.
- **Redefinir la plataforma (transportar y redefinir)**: traslade una aplicación a la nube e introduzca algún nivel de optimización para aprovechar las capacidades de la nube. Ejemplo: Migrar la base de datos Oracle en las instalaciones a Amazon Relational Database Service (Amazon RDS) para Oracle en la nube de Nube de AWS.
- **Recomprar (readquirir)**: cambie a un producto diferente, lo cual se suele llevar a cabo al pasar de una licencia tradicional a un modelo SaaS. Ejemplo: migre su sistema de gestión de relaciones con los clientes (CRM) a Salesforce.com.
- **Volver a alojar (migrar mediante lift-and-shift)**: traslade una aplicación a la nube sin hacer cambios para aprovechar las funcionalidades de la nube. Ejemplo: Migrar la base de datos de Oracle en las instalaciones a Oracle en una instancia de EC2 en la Nube de AWS.
- **Reubicar**: (migrar el hipervisor mediante lift and shift): traslade la infraestructura a la nube sin comprar equipo nuevo, reescribir aplicaciones o modificar las operaciones actuales. Los servidores se migran de una plataforma en las instalaciones a un servicio en la nube para la misma plataforma. Ejemplo: migrar una Microsoft Hyper-V aplicación a AWS.
- **Retener (revisitar)**: conserve las aplicaciones en el entorno de origen. Estas pueden incluir las aplicaciones que requieren una refactorización importante, que desee posponer para más adelante, y las aplicaciones heredadas que desee retener, ya que no hay ninguna justificación empresarial para migrarlas.

- Retirar: retire o elimine las aplicaciones que ya no sean necesarias en un entorno de origen.

A

A2A () Agent-to-Agent

Un protocolo completo para la colaboración entre agentes que facilita la delegación de tareas y la transferencia de estados.

ABAC

Consulte [control de acceso basado en atributos](#).

servicios abstractos

Consulte [servicios administrados](#).

ACID

Consulte [atomicidad, consistencia, aislamiento, durabilidad](#).

migración activa-activa

Método de migración de bases de datos en el que las bases de datos de origen y destino se mantienen sincronizadas (mediante una herramienta de replicación bidireccional o mediante operaciones de escritura doble) y ambas bases de datos gestionan las transacciones de las aplicaciones conectadas durante la migración. Este método permite la migración en lotes pequeños y controlados, en lugar de requerir una transición única. Es más flexible, pero requiere más trabajo que una [migración activa-pasiva](#).

migración activa-pasiva

Método de migración de bases de datos en el que las bases de datos de origen y destino se mantienen sincronizadas, pero solo la de origen gestiona las transacciones de las aplicaciones conectadas, mientras los datos se replican en la de destino. La base de datos de destino no acepta ninguna transacción durante la migración.

Agente

Un sistema de IA que puede razonar, planificar y tomar medidas de forma autónoma utilizando herramientas para alcanzar los objetivos.

Agent Ops

Prácticas operativas para crear, probar, implementar y ejecutar agentes de IA en producción a escala.

función de agregación

Función SQL que actúa en un grupo de filas y calcula un único valor de devolución para el grupo. Entre los ejemplos de funciones de agregación se incluyen SUM y MAX.

IA

Consulte [inteligencia artificial](#).

AIOps

Consulte [operaciones de inteligencia artificial](#)

anonimización

El proceso de eliminar permanentemente la información personal de un conjunto de datos. La anonimización puede ayudar a proteger la privacidad personal. Los datos anonimizados ya no se consideran datos personales.

antipatronos

Una solución que se utiliza con frecuencia para un problema recurrente en el que la solución es contraproducente, ineficaz o menos eficaz que una alternativa.

control de aplicaciones

Enfoque de seguridad que permite usar de manera exclusiva aplicaciones aprobadas para ayudar a proteger un sistema contra el malware.

cartera de aplicaciones

Recopilación de información detallada sobre cada aplicación que utiliza una organización, incluido el costo de creación y mantenimiento de la aplicación y su valor empresarial. Esta información es clave para [el proceso de detección y análisis de la cartera](#) y ayuda a identificar y priorizar las aplicaciones que se van a migrar, modernizar y optimizar.

inteligencia artificial (IA)

El campo de la informática que se dedica al uso de tecnologías informáticas para realizar funciones cognitivas que suelen estar asociadas a los seres humanos, como el aprendizaje, la resolución de problemas y el reconocimiento de patrones. Para más información, consulte [¿Qué es la inteligencia artificial?](#)

operaciones de inteligencia artificial (AIOps)

El proceso de utilizar técnicas de machine learning para resolver problemas operativos, reducir los incidentes operativos y la intervención humana, y mejorar la calidad del servicio. Para obtener más información sobre cómo se utiliza AIOps en la estrategia de migración de AWS, consulte la [Guía de integración de operaciones](#).

cifrado asimétrico

Algoritmo de cifrado que utiliza un par de claves, una clave pública para el cifrado y una clave privada para el descifrado. Puede compartir la clave pública porque no se utiliza para el descifrado, pero el acceso a la clave privada debe estar sumamente restringido.

atomicidad, consistencia, aislamiento, durabilidad (ACID)

Conjunto de propiedades de software que garantizan la validez de los datos y la fiabilidad operativa de una base de datos, incluso en caso de errores, cortes de energía u otros problemas.

control de acceso basado en atributos (ABAC)

La práctica de crear permisos detallados basados en los atributos del usuario, como el departamento, el puesto de trabajo y el nombre del equipo. Para obtener más información, consulte [ABAC AWS en la](#) documentación AWS Identity and Access Management (IAM).

origen de datos fidedigno

Ubicación en la que se almacena la versión principal de los datos, que se considera la fuente de información más fiable. Puede copiar los datos del origen de datos autorizado a otras ubicaciones con el fin de procesarlos o modificarlos, por ejemplo, anonimizarlos, redactarlos o seudonimizarlos.

Zona de disponibilidad

Una ubicación distinta dentro de una Región de AWS que está aislada de los fallos en otras zonas de disponibilidad y que proporciona una conectividad de red económica y de baja latencia a otras zonas de disponibilidad de la misma región.

AWS Marco de adopción de la nube (AWS CAF)

Un marco de directrices y mejores prácticas AWS para ayudar a las organizaciones a desarrollar un plan eficiente y eficaz para migrar con éxito a la nube. AWS CAF organiza la orientación en seis áreas de enfoque denominadas perspectivas: negocios, personas, gobierno, plataforma, seguridad y operaciones. Las perspectivas empresariales, humanas y de gobernanza se centran en las habilidades y los procesos empresariales; las perspectivas de plataforma, seguridad y

operaciones se centran en las habilidades y los procesos técnicos. Por ejemplo, la perspectiva humana se dirige a las partes interesadas que se ocupan de los Recursos Humanos (RR. HH.), las funciones del personal y la administración de las personas. Desde esta perspectiva, AWS CAF proporciona orientación para el desarrollo, la formación y la comunicación de las personas a fin de preparar a la organización para una adopción exitosa de la nube. Para obtener más información, consulte la [Página web de AWS CAF](#) y el [Documento técnico de AWS CAF](#).

AWS Marco de calificación de la carga de trabajo (AWS WQF)

Herramienta que evalúa las cargas de trabajo de migración de bases de datos, recomienda estrategias de migración y proporciona estimaciones de trabajo. AWS WQF se incluye con AWS Schema Conversion Tool (). AWS SCT Analiza los esquemas de bases de datos y los objetos de código, el código de las aplicaciones, las dependencias y las características de rendimiento y proporciona informes de evaluación.

B

bot malicioso

[Bot](#) destinado a causar interrupciones o daños a personas u organizaciones.

BCP

Consulte [planificación de la continuidad del negocio](#).

gráfico de comportamiento

Una vista unificada e interactiva del comportamiento de los recursos y de las interacciones a lo largo del tiempo. Puede utilizar un gráfico de comportamiento con Amazon Detective para examinar los intentos de inicio de sesión fallidos, las llamadas sospechosas a la API y acciones similares. Para obtener más información, consulte [Datos en un gráfico de comportamiento](#) en la documentación de Detective.

sistema big-endian

Un sistema que almacena primero el byte más significativo. Consulte también [endianidad](#).

clasificación binaria

Un proceso que predice un resultado binario (una de las dos clases posibles). Por ejemplo, es posible que su modelo de ML necesite predecir problemas como “¿Este correo electrónico es spam o no es spam?” o “¿Este producto es un libro o un automóvil?”.

filtro de floración

Estructura de datos probabilística y eficiente en términos de memoria que se utiliza para comprobar si un elemento es miembro de un conjunto.

blue/green despliegue

Estrategia de implementación en la que se crean dos entornos separados, pero idénticos. La versión actual de la aplicación se ejecuta en un entorno (azul) y la nueva versión de la aplicación se ejecuta en el otro entorno (verde). Esta estrategia lo ayuda a hacer reversiones rápidas con un impacto mínimo.

bot

Aplicación de software que ejecuta tareas automatizadas a través de Internet y simula la actividad o interacción humana. Algunos bots son útiles o beneficiosos, como los rastreadores web que indexan la información de Internet. Otros bots, conocidos como bots maliciosos, tienen como objetivo causar interrupciones o daños a personas u organizaciones.

botnet

Redes de [bots](#) infectadas por [malware](#) y que están bajo el control de una sola parte, conocida como pastor de bots u operador de bots. Las botnets son el mecanismo más conocido para escalar los bots y su impacto.

branch

Área contenida de un repositorio de código. La primera rama que se crea en un repositorio es la rama principal. Puede crear una rama nueva a partir de una rama existente y, a continuación, desarrollar características o corregir errores en la rama nueva. Una rama que se genera para crear una característica se denomina comúnmente rama de característica. Cuando la característica se encuentra lista para su lanzamiento, se vuelve a combinar la rama de característica con la rama principal. Para obtener más información, consulte [Acerca de las sucursales](#) (GitHub documentación).

acceso de emergencia

En circunstancias excepcionales y mediante un proceso aprobado, es una forma rápida de que un usuario pueda acceder a un Cuenta de AWS sitio al que normalmente no tiene permisos de acceso. Para obtener más información, consulte el indicador de [implementación de procedimientos rompe-cristales](#) en la AWS Well-Architected guía.

estrategia de implementación sobre infraestructura existente

La infraestructura existente en su entorno. Al adoptar una estrategia de implementación sobre infraestructura existente para una arquitectura de sistemas, se diseña la arquitectura en función de las limitaciones de los sistemas y la infraestructura actuales. Si está ampliando la infraestructura existente, puede combinar las estrategias de implementación sobre infraestructuras existentes y de [implementación desde cero](#).

caché de búfer

El área de memoria donde se almacenan los datos a los que se accede con más frecuencia.

capacidad empresarial

Lo que hace una empresa para generar valor (por ejemplo, ventas, servicio al cliente o marketing). Las arquitecturas de microservicios y las decisiones de desarrollo pueden estar impulsadas por las capacidades empresariales. Para obtener más información, consulte la sección [Organizado en torno a las capacidades empresariales](#) del documento técnico [Ejecutar microservicios en contenedores en AWS](#).

planificación de la continuidad del negocio (BCP)

Plan que aborda el posible impacto de un evento disruptivo, como una migración a gran escala en las operaciones y permite a la empresa reanudar las operaciones rápidamente.

C

CAF

Consulte [AWS Cloud Adoption Framework](#).

implementación canario

Lanzamiento lento e incremental de una versión para los usuarios finales. Cuando tenga mayor confianza en la nueva versión, la implementa y reemplaza la versión actual en su totalidad.

CCoE

Consulte [Centro de excelencia en la nube](#).

CDC

Consulte [captura de datos de cambios](#).

captura de datos de cambio (CDC)

Proceso de seguimiento de los cambios en un origen de datos, como una tabla de base de datos, y registro de los metadatos relacionados con el cambio. Puede utilizar los CDC para diversos fines, como auditar o replicar los cambios en un sistema de destino para mantener la sincronización.

ingeniería del caos

Introducción intencionada de fallos o eventos disruptivos para poner a prueba la resiliencia de un sistema. Puedes usar [AWS Fault Injection Service \(AWS FIS\)](#) para realizar experimentos que estresen tus AWS cargas de trabajo y evalúen su respuesta.

CI/CD

Consulte [integración continua y entrega continua](#).

clasificación

Un proceso de categorización que permite generar predicciones. Los modelos de ML para problemas de clasificación predicen un valor discreto. Los valores discretos siempre son distintos entre sí. Por ejemplo, es posible que un modelo necesite evaluar si hay o no un automóvil en una imagen.

Desarrollador ciudadano

Un usuario empresarial que crea aplicaciones de IA utilizando plataformas sin code/low código sin conocimientos técnicos especializados.

cifrado del cliente

Cifrado de datos localmente, antes de que el objetivo los Servicio de AWS reciba.

Centro de excelencia en la nube (CCoE)

Equipo multidisciplinario que impulsa los esfuerzos de adopción de la nube en toda la organización, incluido el desarrollo de las prácticas recomendadas en la nube, la movilización de recursos, el establecimiento de plazos de migración y la dirección de la organización durante las transformaciones a gran escala. Para obtener más información, consulte las [publicaciones de CCoE](#) en el blog de estrategia Nube de AWS empresarial.

computación en la nube

La tecnología en la nube que se utiliza normalmente para la administración de dispositivos de IoT y el almacenamiento de datos de forma remota. La computación en la nube suele estar relacionada con la tecnología de [computación de periferia](#).

modelo operativo en la nube

En una organización de TI, el modelo operativo que se utiliza para crear, madurar y optimizar uno o más entornos de nube. Para obtener más información, consulte [Creación de su modelo operativo de nube](#).

etapas de adopción de la nube

Las siguientes son las cuatro fases por las que suelen pasar las empresas cuando migran a la Nube de AWS:

- Proyecto: ejecución de algunos proyectos relacionados con la nube con fines de prueba de concepto y aprendizaje
- Fundamento: realización de inversiones fundamentales para escalar la adopción de la nube (p. ej., crear una zona de aterrizaje, definir un CCoE, establecer un modelo de operaciones)
- Migración: migración de aplicaciones individuales
- Re-invention — Optimizar los productos y servicios e innovar en la nube

Stephen Orban definió estas etapas en la entrada del blog The [Journey Toward Cloud-First & the Stages of Adoption del](#) blog Nube de AWS Enterprise Strategy. Para obtener información sobre su relación con la estrategia de AWS migración, consulte la [guía de preparación para la migración](#).

CMDB

Consulte [base de datos de administración de configuración](#).

repositorio de código

Una ubicación donde el código fuente y otros activos, como documentación, muestras y scripts, se almacenan y actualizan mediante procesos de control de versiones. Algunos repositorios en la nube comunes son GitHub o Bitbucket Cloud. Cada versión del código se denomina rama. En una estructura de microservicios, cada repositorio se encuentra dedicado a una única funcionalidad. Una sola CI/CD canalización puede utilizar varios repositorios.

caché en frío

Una caché de búfer que está vacía no está bien poblada o contiene datos obsoletos o irrelevantes. Esto afecta al rendimiento, ya que la instancia de la base de datos debe leer desde la memoria principal o el disco, lo que es más lento que leer desde la memoria caché del búfer.

datos fríos

Datos a los que se accede con poca frecuencia y que suelen ser históricos. Al consultar este tipo de datos, normalmente se aceptan consultas lentas. Trasladar estos datos a niveles o clases de almacenamiento de menor rendimiento y menos costosos puede reducir los costos.

visión artificial (CV)

Campo de la [IA](#) que utiliza el machine learning para analizar y extraer información de formatos visuales, como imágenes y videos digitales. Por ejemplo, Amazon SageMaker AI proporciona algoritmos de procesamiento de imágenes para CV.

deriva de configuración

En el caso de una carga de trabajo, un cambio en la configuración con respecto al estado esperado. Podría provocar que la carga de trabajo deje de cumplir las normas y, por lo general, es gradual e involuntaria.

base de datos de administración de configuración (CMDB)

Repositorio que almacena y administra información sobre una base de datos y su entorno de TI, incluidos los componentes de hardware y software y sus configuraciones. Por lo general, los datos de una CMDB se utilizan en la etapa de detección y análisis de la cartera de productos durante la migración.

paquete de conformidad

Un conjunto de AWS Config reglas y medidas correctivas que puede reunir para personalizar sus controles de conformidad y seguridad. Puede implementar un paquete de conformidad como una entidad única en una región Cuenta de AWS y, o en una organización, mediante una plantilla YAML. Para obtener más información, consulta los [paquetes de conformidad](#) en la documentación. AWS Config

integración y entrega continuas (I) CI/CD

El proceso de automatización de las etapas de origen, creación, prueba, puesta en escena y producción del proceso de publicación del software. CI/CD se describe comúnmente como una canalización. CI/CD puede ayudarlo a automatizar los procesos, mejorar la productividad, mejorar la calidad del código y entregar más rápido. Para obtener más información, consulte [Beneficios de la entrega continua](#). CD también puede significar implementación continua. Para obtener más información, consulte [Entrega continua frente a implementación continua](#).

CV

Consulte [visión artificial](#).

D

datos en reposo

Datos que están estacionarios en la red, como los datos que se encuentran almacenados.

clasificación de datos

Un proceso para identificar y clasificar los datos de su red en función de su importancia y sensibilidad. Es un componente fundamental de cualquier estrategia de administración de riesgos de ciberseguridad porque lo ayuda a determinar los controles de protección y retención adecuados para los datos. La clasificación de los datos es un componente del pilar de seguridad del AWS Well-Architected Framework. Para obtener más información, consulte [Clasificación de datos](#).

deriva de datos

Una variación significativa entre los datos de producción y los datos que se utilizaron para entrenar un modelo de machine learning, o un cambio significativo en los datos de entrada a lo largo del tiempo. La deriva de datos puede reducir la calidad, la precisión y la imparcialidad generales de las predicciones de los modelos de machine learning.

datos en tránsito

Datos que se mueven de forma activa por la red, por ejemplo, entre los recursos de la red.

mallado de datos

Marco de arquitectura que proporciona una propiedad de datos distribuida y descentralizada con una administración y una gobernanza centralizadas.

minimización de datos

El principio de recopilar y procesar solo los datos estrictamente necesarios. Practicar la minimización de los datos Nube de AWS puede reducir los riesgos de privacidad, los costos y la huella de carbono de la analítica.

perímetro de datos

Un conjunto de barreras preventivas en su AWS entorno que ayudan a garantizar que solo las identidades confiables accedan a los recursos confiables desde las redes esperadas. Para obtener más información, consulte [Crear un perímetro de datos sobre](#) AWS

preprocesamiento de datos

Transformar los datos sin procesar en un formato que su modelo de ML pueda analizar fácilmente. El preprocesamiento de datos puede implicar eliminar determinadas columnas o filas y corregir los valores faltantes, incoherentes o duplicados.

procedencia de los datos

El proceso de rastrear el origen y el historial de los datos a lo largo de su ciclo de vida, por ejemplo, la forma en que se generaron, transmitieron y almacenaron los datos.

titular de los datos

Persona cuyos datos se recopilan y procesan.

almacenamiento de datos

Sistema de administración de datos que respalda la inteligencia empresarial, como los análisis. Los almacenes de datos suelen contener grandes cantidades de datos históricos y, por lo general, se utilizan para las consultas y los análisis.

lenguaje de definición de datos (DDL)

Instrucciones o comandos para crear o modificar la estructura de tablas y objetos de una base de datos.

lenguaje de manipulación de datos (DML)

Instrucciones o comandos para modificar (insertar, actualizar y eliminar) la información de una base de datos.

DDL

Consulte [lenguaje de definición de bases de datos](#).

conjunto profundo

Combinar varios modelos de aprendizaje profundo para la predicción. Puede utilizar conjuntos profundos para obtener una predicción más precisa o para estimar la incertidumbre de las predicciones.

aprendizaje profundo

Un subcampo del ML que utiliza múltiples capas de redes neuronales artificiales para identificar el mapeo entre los datos de entrada y las variables objetivo de interés.

defensa en profundidad

Un enfoque de seguridad de la información en el que se distribuyen cuidadosamente una serie de mecanismos y controles de seguridad en una red informática para proteger la confidencialidad, la integridad y la disponibilidad de la red y de los datos que contiene. Al adoptar esta estrategia AWS, se añaden varios controles en diferentes capas de la AWS Organizations estructura para ayudar a proteger los recursos. Por ejemplo, un enfoque de defensa en profundidad podría combinar la autenticación multifactor, la segmentación de la red y el cifrado.

administrador delegado

En AWS Organizations, un servicio compatible puede registrar una cuenta de AWS miembro para administrar las cuentas de la organización y gestionar los permisos de ese servicio. Esta cuenta se denomina administrador delegado para ese servicio. Para obtener más información y una lista de servicios compatibles, consulte [Servicios que funcionan con AWS Organizations](#) en la documentación de AWS Organizations .

Implementación

El proceso de hacer que una aplicación, características nuevas o correcciones de código se encuentren disponibles en el entorno de destino. La implementación abarca implementar cambios en una base de código y, a continuación, crear y ejecutar esa base en los entornos de la aplicación.

entorno de desarrollo

Consulte [entorno](#).

control de detección

Un control de seguridad que se ha diseñado para detectar, registrar y alertar después de que se produzca un evento. Estos controles son una segunda línea de defensa, ya que lo advierten sobre los eventos de seguridad que han eludido los controles preventivos establecidos. Para obtener más información, consulte [Controles de detección](#) en Implementación de controles de seguridad en AWS.

asignación de flujos de valor para el desarrollo (DVSM)

Proceso que se utiliza para identificar y priorizar las restricciones que afectan negativamente a la velocidad y la calidad en el ciclo de vida del desarrollo de software. DVSM amplía el proceso de asignación del flujo de valor diseñado originalmente para las prácticas de fabricación ajustada. Se centra en los pasos y los equipos necesarios para crear y transferir valor a través del proceso de desarrollo de software.

gemelo digital

Representación virtual de un sistema del mundo real, como un edificio, una fábrica, un equipo industrial o una línea de producción. Los gemelos digitales son compatibles con el mantenimiento predictivo, la supervisión remota y la optimización de la producción.

tabla de dimensiones

En un [esquema en estrella](#), tabla más pequeña que contiene los atributos de datos sobre los datos cuantitativos en una tabla de hechos. Los atributos de la tabla de dimensiones suelen ser campos de texto o números discretos que se comportan como texto. Estos atributos se suelen utilizar para restringir consultas, filtrarlas y etiquetar los conjuntos de resultados.

desastre

Un evento que impide que una carga de trabajo o un sistema cumplan sus objetivos empresariales en su ubicación principal de implementación. Estos eventos pueden ser desastres naturales, fallos técnicos o el resultado de acciones humanas, como una configuración incorrecta involuntaria o un ataque de malware.

recuperación de desastres (DR)

Estrategia y proceso que utiliza para minimizar el tiempo de inactividad y la pérdida de datos a causa de un [desastre](#). Para obtener más información, consulte [Recuperación de cargas de trabajo ante desastres en AWS: Recuperación en la nube](#) en el AWS Well-Architected marco.

DML

Consulte [lenguaje de manipulación de bases de datos](#).

diseño basado en el dominio

Un enfoque para desarrollar un sistema de software complejo mediante la conexión de sus componentes a dominios en evolución, o a los objetivos empresariales principales, a los que sirve cada componente. Eric Evans introdujo este concepto en su libro *Domain-Driven Design: Tackling Complexity in the Heart of Software* (Boston: Addison-Wesley Professional, 2003). Para obtener información sobre cómo utilizar el diseño basado en dominios con el patrón de higos estranguladores, consulte [Modernización gradual de los servicios web antiguos de ASP.NET Microsoft \(ASMX\) mediante contenedores y Amazon API Gateway](#).

DR

Consulte [recuperación ante desastres](#).

Detección de desviaciones

Seguimiento de las desviaciones con respecto a una configuración con línea de base. Por ejemplo, puedes usarlo AWS CloudFormation para [detectar desviaciones en los recursos del sistema](#) o puedes usarlo AWS Control Tower para [detectar cambios en tu landing zone](#) que puedan afectar al cumplimiento de los requisitos de gobierno.

DVSM

Consulte [asignación de flujos de valor para el desarrollo](#).

E

EDA

Consulte [análisis de datos de tipo exploratorio](#).

EDI

Consulte [intercambio electrónico de datos](#).

computación en la periferia

La tecnología que aumenta la potencia de cálculo de los dispositivos inteligentes en la periferia de una red de IoT. En comparación con la [computación en la nube](#), la computación de periferia puede reducir la latencia de la comunicación y mejorar el tiempo de respuesta.

intercambio electrónico de datos (EDI)

Intercambio automatizado de documentos comerciales entre organizaciones. Para más información, consulte [¿Qué es el intercambio electrónico de datos?](#)

cifrado

Proceso de computación que transforma datos de texto plano, que son legibles por humanos, en texto cifrado.

clave de cifrado

Cadena criptográfica de bits aleatorios que se genera mediante un algoritmo de cifrado. Las claves pueden variar en longitud y cada una se ha diseñado para ser impredecible y única.

endianidad

El orden en el que se almacenan los bytes en la memoria del ordenador. Big-endian los sistemas almacenan primero el byte más significativo. Little-endian los sistemas almacenan primero el byte menos significativo.

punto de conexión

Consulte [punto de conexión de servicio](#).

servicio de punto de conexión

Servicio que puede alojar en una nube privada virtual (VPC) para compartir con otros usuarios. Puede crear un servicio de punto final con AWS PrivateLink entidades principales Cuentas de AWS o AWS Identity and Access Management (de IAM) y conceder permisos a ellas. Estas cuentas o entidades principales pueden conectarse a su servicio de punto de conexión de forma privada mediante la creación de puntos de conexión de VPC de interfaz. Para obtener más información, consulte [Creación de un servicio de punto de conexión](#) en la documentación de Amazon Virtual Private Cloud (Amazon VPC).

planificación de recursos empresariales (ERP)

Sistema que automatiza y administra los procesos empresariales clave (como la contabilidad, [MES](#) y la administración de proyectos) de una empresa.

cifrado de sobre

El proceso de cifrar una clave de cifrado con otra clave de cifrado. Para obtener más información, consulte el [cifrado de sobres](#) en la documentación de AWS Key Management Service (AWS KMS).

entorno

Una instancia de una aplicación en ejecución. Los siguientes son los tipos de entornos más comunes en la computación en la nube:

- entorno de desarrollo: instancia de una aplicación en ejecución que solo se encuentra disponible para el equipo principal responsable del mantenimiento de la aplicación. Los entornos de desarrollo se utilizan para probar los cambios antes de promocionarlos a los entornos superiores. Este tipo de entorno a veces se denomina entorno de prueba.
- entornos inferiores: todos los entornos de desarrollo de una aplicación, como los que se utilizan para las compilaciones y pruebas iniciales.

- entorno de producción: instancia de una aplicación en ejecución a la que pueden acceder los usuarios finales. En un CI/CD proceso, el entorno de producción es el último entorno de implementación.
- entornos superiores: todos los entornos a los que pueden acceder usuarios que no sean del equipo de desarrollo principal. Esto puede incluir un entorno de producción, entornos de preproducción y entornos para las pruebas de aceptación por parte de los usuarios.

epopeya

En las metodologías ágiles, son categorías funcionales que ayudan a organizar y priorizar el trabajo. Las epopeyas brindan una descripción detallada de los requisitos y las tareas de implementación. Por ejemplo, las epopeyas AWS de seguridad de CAF incluyen la gestión de identidades y accesos, los controles de detección, la seguridad de la infraestructura, la protección de datos y la respuesta a incidentes. Para obtener más información sobre las epopeyas en la estrategia de migración de AWS , consulte la [Guía de implementación del programa](#).

ERP

Consulte [planificación de recursos empresariales](#).

análisis de datos de tipo exploratorio (EDA)

El proceso de analizar un conjunto de datos para comprender sus características principales. Se recopilan o agregan datos y, a continuación, se realizan las investigaciones iniciales para encontrar patrones, detectar anomalías y comprobar las suposiciones. El EDA se realiza mediante el cálculo de estadísticas resumidas y la creación de visualizaciones de datos.

F

tabla de hechos

Tabla central de un [esquema en estrella](#). Almacena datos cuantitativos sobre operaciones empresariales. Por lo general, una tabla de hechos contiene dos tipos de columnas: las que contienen medidas y las que contienen una clave externa para una tabla de dimensiones.

Fail Fast

Filosofía que utiliza pruebas frecuentes e incrementales para reducir el ciclo de vida del desarrollo. Es una parte fundamental de los enfoques ágiles.

límite de aislamiento de errores

En el Nube de AWS, un límite, como una zona de disponibilidad Región de AWS, un plano de control o un plano de datos, que limita el efecto de una falla y ayuda a mejorar la resiliencia de las cargas de trabajo. Para más información, consulte [AWS Fault Isolation Boundaries](#).

rama de característica

Consulte [rama](#).

características

Los datos de entrada que se utilizan para hacer una predicción. Por ejemplo, en un contexto de fabricación, las características pueden ser imágenes que se capturan periódicamente desde la línea de fabricación.

importancia de las características

La importancia que tiene una característica para las predicciones de un modelo. Por lo general, esto se expresa como una puntuación numérica que se puede calcular mediante diversas técnicas, como las explicaciones aditivas de Shapley (SHAP) y los gradientes integrados. Para obtener más información, consulte [Interpretabilidad del modelo de aprendizaje automático](#) con AWS.

transformación de funciones

Optimizar los datos para el proceso de ML, lo que incluye enriquecer los datos con fuentes adicionales, escalar los valores o extraer varios conjuntos de información de un solo campo de datos. Esto permite que el modelo de ML se beneficie de los datos. Por ejemplo, si divide la fecha del “27 de mayo de 2021 00:15:37” en “jueves”, “mayo”, “2021” y “15”, puede ayudar al algoritmo de aprendizaje a aprender patrones matizados asociados a los diferentes componentes de los datos.

peticiones con pocos pasos

Proporcionar a un [LLM](#) una pequeña cantidad de ejemplos que demuestren la tarea y el resultado deseado antes de pedirle que lleve a cabo una tarea similar. Esta técnica es una aplicación del aprendizaje contextual, en el que los modelos aprenden a partir de ejemplos (tomas) integrados en las instrucciones. Few-shot Las indicaciones pueden ser eficaces para tareas que requieren un formato, un razonamiento o un conocimiento del dominio específicos. Consulte también [peticiones desde cero](#).

FGAC

Consulte [control de acceso detallado](#).

control de acceso preciso (FGAC)

El uso de varias condiciones que tienen por objetivo permitir o denegar una solicitud de acceso.

migración relámpago

Método de migración de bases de datos que utiliza la replicación continua de datos mediante la [captura de datos de cambio](#) para migrar los datos en el menor tiempo posible, en lugar de utilizar un enfoque gradual. El objetivo es reducir al mínimo el tiempo de inactividad.

FM

Consulte [modelo fundacional](#).

Modelo fundacional (FM)

Gran red neuronal de aprendizaje profundo que se entrenó con conjuntos de datos masivos de datos generalizados y no etiquetados. Los FM pueden hacer una amplia variedad de tareas generales, como comprender el lenguaje, generar texto e imágenes y conversar en lenguaje natural. Para más información, consulte [¿Qué son los modelos fundacionales?](#)

Puerta de enlace FM

Un intermediario centralizado que controla y normaliza el acceso a los modelos básicos. También se conoce como puerta de enlace LLM.

G

IA generativa

Subconjunto de modelos de [IA](#) que se entrenaron con grandes cantidades de datos y que pueden utilizar una simple petición de texto para crear contenido y artefactos nuevos, como imágenes, videos, texto y audio. Para más información, consulte [¿Qué es la IA generativa?](#)

bloqueo geográfico

Consulte [restricciones geográficas](#).

restricciones geográficas (bloqueo geográfico)

En Amazon CloudFront, una opción para impedir que los usuarios de países específicos accedan a las distribuciones de contenido. Puede utilizar una lista de permitidos o bloqueados para especificar los países aprobados y prohibidos. Para obtener más información, consulta [Restringir la distribución geográfica del contenido](#) en la CloudFront documentación.

Flujo de trabajo de Gitflow

Un enfoque en el que los entornos inferiores y superiores utilizan diferentes ramas en un repositorio de código fuente. El flujo de trabajo de Gitflow se considera heredado, mientras que el [flujo de trabajo basado en enlaces troncales](#) es el enfoque moderno preferido.

imagen dorada

Instantánea de un sistema o software que se usa como plantilla para implementar nuevas instancias de ese sistema o software. Por ejemplo, en la fabricación, una imagen dorada se puede utilizar para aprovisionar software en varios dispositivos y ayuda a mejorar la velocidad, la escalabilidad y la productividad de las operaciones de fabricación de dispositivos.

estrategia de implementación desde cero

La ausencia de infraestructura existente en un entorno nuevo. Al adoptar una estrategia de implementación desde cero para una arquitectura de sistemas, puede seleccionar todas las tecnologías nuevas sin que estas deban ser compatibles con una infraestructura existente, lo que también se conoce como [implementación sobre infraestructura existente](#). Si está ampliando la infraestructura existente, puede combinar las estrategias de implementación sobre infraestructuras existentes y de implementación desde cero.

barrera de protección

Una regla de alto nivel que ayuda a regular los recursos, las políticas y la conformidad en todas las unidades organizativas (OU). Las barreras de protección preventivas aplican políticas para garantizar la alineación con los estándares de conformidad. Se implementan mediante políticas de control de servicios y límites de permisos de IAM. Las barreras de protección de detección detectan las vulneraciones de las políticas y los problemas de conformidad, y generan alertas para su corrección. Se implementan mediante Amazon AWS Config AWS Security Hub CSPM GuardDuty AWS Trusted Advisor, Amazon Inspector y AWS Lambda cheques personalizados.

barandas (AI)

Mecanismos de seguridad que filtran, validan y restringen las entradas y salidas de los [agentes](#) para ayudar a garantizar un comportamiento responsable y seguro de la IA.

H

HA

Consulte [alta disponibilidad](#).

migración heterogénea de bases de datos

Migración de la base de datos de origen a una base de datos de destino que utilice un motor de base de datos diferente (por ejemplo, de Oracle a Amazon Aurora). La migración heterogénea suele ser parte de un esfuerzo de rediseño de la arquitectura y convertir el esquema puede ser una tarea compleja. [AWS ofrece AWS SCT](#), lo cual ayuda con las conversiones de esquemas.

alta disponibilidad (HA)

La capacidad de una carga de trabajo para funcionar de forma continua, sin intervención, en caso de desafíos o desastres. Los sistemas de alta disponibilidad están diseñados para realizar una conmutación por error automática, ofrecer un rendimiento de alta calidad de forma constante y gestionar diferentes cargas y fallos con un impacto mínimo en el rendimiento.

modernización histórica

Un enfoque utilizado para modernizar y actualizar los sistemas de tecnología operativa (TO) a fin de satisfacer mejor las necesidades de la industria manufacturera. Un histórico es un tipo de base de datos que se utiliza para recopilar y almacenar datos de diversas fuentes en una fábrica.

datos de reserva

Parte de los datos históricos etiquetados que se ocultan de un conjunto de datos que se utiliza para entrenar un modelo de [machine learning](#). Puede utilizar los datos de reserva para evaluar el rendimiento del modelo mediante la comparación de las predicciones del modelo con los datos de reserva.

human-in-the-loop (HiTL)

Un patrón de flujo de trabajo en el que la ejecución de los [agentes](#) se detiene para su revisión y aprobación por parte de una persona en los puntos de decisión críticos.

migración homogénea de bases de datos

Migración de la base de datos de origen a una base de datos de destino que comparte el mismo motor de base de datos (por ejemplo, Microsoft SQL Server a Amazon RDS para SQL Server). La migración homogénea suele formar parte de un esfuerzo para volver a alojar o redefinir la plataforma. Puede utilizar las utilidades de bases de datos nativas para migrar el esquema.

datos recientes

Datos a los que se accede con frecuencia, como datos en tiempo real o datos traslacionales recientes. Por lo general, estos datos requieren un nivel o una clase de almacenamiento de alto rendimiento para proporcionar respuestas rápidas a las consultas.

hotfix

Una solución urgente para un problema crítico en un entorno de producción. Debido a su urgencia, una revisión suele realizarse fuera del flujo de trabajo habitual de las DevOps versiones.

periodo de hiperatención

Periodo, inmediatamente después de la transición, durante el cual un equipo de migración administra y monitorea las aplicaciones migradas en la nube para solucionar cualquier problema. Por lo general, este periodo dura de 1 a 4 días. Al final del periodo de hiperatención, el equipo de migración suele transferir la responsabilidad de las aplicaciones al equipo de operaciones en la nube.

I

laC

Consulte [infraestructura como código](#).

políticas basadas en identidades

Política asociada a uno o más directores de IAM que define sus permisos en el entorno. Nube de AWS

aplicación inactiva

Aplicación que utiliza un promedio de CPU y memoria de entre 5 y 20 por ciento durante un periodo de 90 días. En un proyecto de migración, es habitual retirar estas aplicaciones o mantenerlas en las instalaciones.

IIoT

Consulte [Internet de las cosas industrial](#).

infraestructura inmutable

Modelo que implementa una nueva infraestructura para las cargas de trabajo de producción en lugar de actualizar o modificar la infraestructura existente o aplicarle revisiones. Las infraestructuras inmutables son de manera intrínseca más coherentes, fiables y predecibles que las [infraestructuras mutables](#). Para obtener más información, consulte las mejores prácticas del [Framework para implementar con una infraestructura inmutable](#). AWS Well-Architected

VPC entrante (de entrada)

En una arquitectura de AWS cuentas múltiples, una VPC que acepta, inspecciona y enruta las conexiones de red desde fuera de una aplicación. La [Arquitectura de referencia de seguridad de AWS](#) recomienda configurar su cuenta de red con VPC entrantes, salientes y de inspección para proteger la interfaz bidireccional entre su aplicación e Internet en general.

migración gradual

Estrategia de transición en la que se migra la aplicación en partes pequeñas en lugar de realizar una transición única y completa. Por ejemplo, puede trasladar inicialmente solo unos pocos microservicios o usuarios al nuevo sistema. Tras comprobar que todo funciona correctamente, puede trasladar microservicios o usuarios adicionales de forma gradual hasta que pueda retirar su sistema heredado. Esta estrategia reduce los riesgos asociados a las grandes migraciones.

Industria 4.0

Un término que [Klaus Schwab](#) introdujo en 2016 para referirse a la modernización de los procesos de fabricación mediante avances en la conectividad, los datos en tiempo real, la automatización, el análisis y. AI/ML

infraestructura

Todos los recursos y activos que se encuentran en el entorno de una aplicación.

infraestructura como código (IaC)

Proceso de aprovisionamiento y administración de la infraestructura de una aplicación mediante un conjunto de archivos de configuración. La IaC se ha diseñado para ayudarlo a centralizar la administración de la infraestructura, estandarizar los recursos y escalar con rapidez a fin de que los entornos nuevos sean repetibles, fiables y consistentes.

Internet de las cosas industrial (IIoT)

El uso de sensores y dispositivos conectados a Internet en los sectores industriales, como el productivo, el eléctrico, el automotriz, el sanitario, el de las ciencias de la vida y el de la agricultura. Para obtener más información, consulte [Creación de una estrategia de transformación digital del Internet de las cosas industrial \(IIoT\)](#).

VPC de inspección

En una arquitectura de AWS cuentas múltiples, una VPC centralizada que gestiona las inspecciones del tráfico de red entre las VPC (iguales o Regiones de AWS diferentes), Internet y las redes locales. La [Arquitectura de referencia de seguridad de AWS](#) recomienda configurar su

cuenta de red con VPC entrantes, salientes y de inspección para proteger la interfaz bidireccional entre su aplicación e Internet en general.

Internet de las cosas (IoT)

Red de objetos físicos conectados con sensores o procesadores integrados que se comunican con otros dispositivos y sistemas a través de Internet o de una red de comunicación local. Para obtener más información, consulte [¿Qué es IoT?](#).

interpretabilidad

Característica de un modelo de machine learning que describe el grado en que un ser humano puede entender cómo las predicciones del modelo dependen de sus entradas. Para obtener más información, consulte Interpretabilidad del modelo [de aprendizaje automático](#) con AWS

IoT

Consulte [Internet de las cosas](#).

biblioteca de información de TI (ITIL)

Conjunto de prácticas recomendadas para ofrecer servicios de TI y alinearlos con los requisitos empresariales. La ITIL proporciona la base para la ITSM.

administración de servicios de TI (ITSM)

Actividades asociadas con el diseño, la implementación, la administración y el soporte de los servicios de TI para una organización. Para obtener información sobre la integración de las operaciones en la nube con las herramientas de ITSM, consulte la [Guía de integración de operaciones](#).

ITIL

Consulte [biblioteca de información de TI](#).

ITSM

Consulte [administración de servicios de TI](#).

L

control de acceso basado en etiquetas (LBAC)

Una implementación del control de acceso obligatorio (MAC) en la que a los usuarios y a los propios datos se les asigna explícitamente un valor de etiqueta de seguridad. La intersección

entre la etiqueta de seguridad del usuario y la etiqueta de seguridad de los datos determina qué filas y columnas puede ver el usuario.

zona de aterrizaje

Una landing zone es un AWS entorno multicuenta bien diseñado, escalable y seguro. Este es un punto de partida desde el cual las empresas pueden lanzar e implementar rápidamente cargas de trabajo y aplicaciones con confianza en su entorno de seguridad e infraestructura. Para obtener más información sobre las zonas de aterrizaje, consulte [Configuración de un entorno de AWS seguro y escalable con varias cuentas](#).

modelo de lenguaje de gran tamaño (LLM)

Modelo de [IA](#) de aprendizaje profundo que se entrenó previamente con una gran cantidad de datos. Un LLM puede llevar a cabo varias tareas, como responder preguntas, resumir documentos, traducir textos a otros idiomas y completar oraciones. Para más información, consulte [¿Qué es un LLM \(modelo de lenguaje de gran tamaño\)?](#)

migración grande

Migración de 300 servidores o más.

LBAC

Consulte [control de acceso basado en etiquetas](#).

privilegio mínimo

La práctica recomendada de seguridad que consiste en conceder los permisos mínimos necesarios para realizar una tarea. Para obtener más información, consulte [Aplicar permisos de privilegio mínimo](#) en la documentación de IAM.

migrar mediante lift-and-shift

Consulte [Las 7 R](#).

sistema little-endian

Un sistema que almacena primero el byte menos significativo. Consulte también [endianidad](#).

LLM

Consulte [modelo de lenguaje de gran tamaño](#).

entornos inferiores

Consulte [entorno](#).

M

machine learning (ML)

Un tipo de inteligencia artificial que utiliza algoritmos y técnicas para el reconocimiento y el aprendizaje de patrones. El ML analiza y aprende de los datos registrados, como los datos del Internet de las cosas (IoT), para generar un modelo estadístico basado en patrones. Para más información, consulte [Machine learning](#).

rama principal

Consulte [rama](#).

malware

Software diseñado para comprometer la seguridad o la privacidad de la computadora. El malware podría interrumpir los sistemas informáticos, filtrar información confidencial u obtener acceso no autorizado. Algunos ejemplos de malware son los virus, los gusanos, el ransomware, los troyanos, el spyware y los registradores de pulsaciones de teclas.

Servicios administrados

Servicios de AWS en el que AWS opera la capa de infraestructura, el sistema operativo y las plataformas, y se accede a los puntos finales para almacenar y recuperar datos. Amazon Simple Storage Service (Amazon S3) y Amazon DynamoDB son ejemplos de servicios administrados. También se conocen como servicios abstractos.

sistema de ejecución de fabricación (MES)

Sistema de software para seguir, supervisar, documentar y controlar los procesos de producción que convierten las materias primas en productos acabados en la zona de producción.

MAP

Consulte [Programa de aceleración de la migración](#).

MCP

Consulte [Model Context Protocol](#).

Protocolo de contexto para modelos (MCP)

Un protocolo sin estado para la comunicación entre el [agente](#) y la [herramienta](#).

Servidor MCP

Un servicio que expone una o más [herramientas](#) a través del protocolo [Model Context](#).

mecanismo

Proceso completo mediante el que se crea una herramienta, se impulsa su adopción y, a continuación, se inspeccionan los resultados para hacer ajustes. Un mecanismo es un ciclo que se refuerza y mejora por sí mismo a medida que funciona. Para obtener más información, consulte [Creación de mecanismos](#) en el AWS Well-Architected marco.

cuenta de miembro

Todas las Cuentas de AWS demás cuentas, excepto la de administración, que forman parte de una organización AWS Organizations. Una cuenta no puede pertenecer a más de una organización a la vez.

MES

Consulte [sistema de ejecución de fabricación](#).

Message Queuing Telemetry Transport (MQTT)

[Un protocolo de comunicación ligero de máquina a máquina \(M2M\), basado en el publish/subscribe patrón, para dispositivos de IoT con recursos limitados.](#)

microservicio

Un servicio pequeño e independiente que se comunica a través de API bien definidas y que, por lo general, es propiedad de equipos pequeños e independientes. Por ejemplo, un sistema de seguros puede incluir microservicios que se adapten a las capacidades empresariales, como las de ventas o marketing, o a subdominios, como las de compras, reclamaciones o análisis. Los beneficios de los microservicios incluyen la agilidad, la escalabilidad flexible, la facilidad de implementación, el código reutilizable y la resiliencia. Para obtener más información, consulte [Integrar](#) microservicios mediante servicios sin servidor. AWS

arquitectura de microservicios

Un enfoque para crear una aplicación con componentes independientes que ejecutan cada proceso de la aplicación como un microservicio. Estos microservicios se comunican a través de una interfaz bien definida mediante API ligeras. Cada microservicio de esta arquitectura se puede actualizar, implementar y escalar para satisfacer la demanda de funciones específicas de una aplicación. Para obtener más información, consulte [Implementación de microservicios](#) en. AWS

Programa de aceleración de la migración (MAP)

Un AWS programa que proporciona soporte de consultoría, formación y servicios para ayudar a las organizaciones a crear una base operativa sólida para migrar a la nube y para ayudar a

compensar el costo inicial de las migraciones. El MAP incluye una metodología de migración para ejecutar las migraciones antiguas de forma metódica y un conjunto de herramientas para automatizar y acelerar los escenarios de migración más comunes.

migración a escala

Proceso de transferencia de la mayoría de la cartera de aplicaciones a la nube en oleadas, con más aplicaciones desplazadas a un ritmo más rápido en cada oleada. En esta fase, se utilizan las prácticas recomendadas y las lecciones aprendidas en las fases anteriores para implementar una fábrica de migración de equipos, herramientas y procesos con el fin de agilizar la migración de las cargas de trabajo mediante la automatización y la entrega ágil. Esta es la tercera fase de la [estrategia de migración de AWS](#).

fábrica de migración

Cross-functional equipos que agilizan la migración de las cargas de trabajo mediante enfoques ágiles y automatizados. Los equipos de las fábricas de migración suelen estar compuestos por analistas y propietarios de operaciones, ingenieros de migración, desarrolladores y DevOps profesionales que trabajan a pasos agigantados. Entre el 20 y el 50 por ciento de la cartera de aplicaciones empresariales se compone de patrones repetidos que pueden optimizarse mediante un enfoque de fábrica. Para obtener más información, consulte la [discusión sobre las fábricas de migración](#) y la [Guía de fábricas de migración a la nube](#) en este contenido.

metadatos de migración

Información sobre la aplicación y el servidor que se necesita para completar la migración. Cada patrón de migración requiere un conjunto diferente de metadatos de migración. Algunos ejemplos de metadatos de migración son la subred de destino, el grupo de seguridad y AWS la cuenta.

patrón de migración

Tarea de migración repetible que detalla la estrategia de migración, el destino de la migración y la aplicación o el servicio de migración utilizados. Ejemplo: rehospede la migración a Amazon EC2 AWS con Application Migration Service.

Migration Portfolio Assessment (MPA)

Herramienta en línea que proporciona información a fin de validar los argumentos comerciales necesarios para migrar a la Nube de AWS. La MPA ofrece una evaluación detallada de la cartera (adecuación del tamaño de los servidores, precios, comparaciones del costo total de propiedad, análisis de los costos de migración), así como una planificación de la migración (análisis y recopilación de datos de aplicaciones, agrupación de aplicaciones, priorización de la migración y

planificación de oleadas). La [herramienta MPA](#) (requiere iniciar sesión) está disponible de forma gratuita para todos los AWS consultores y consultores de los socios de APN.

Evaluación de la preparación para la migración (MRA)

Proceso que consiste en obtener información sobre el estado de preparación de una organización para la nube, identificar sus puntos fuertes y débiles y elaborar un plan de acción para cerrar las brechas identificadas mediante el AWS CAF. Para obtener más información, consulte la [Guía de preparación para la migración](#). La MRA es la primera fase de la [estrategia de migración de AWS](#).

estrategia de migración

Enfoque utilizado para migrar una carga de trabajo a la Nube de AWS. Para más información, consulte la entrada [Las 7 R](#) de este glosario y también [Mobilize your organization to accelerate large-scale migrations](#).

ML

Consulte [machine learning](#).

modernización

Transformar una aplicación obsoleta (antigua o monolítica) y su infraestructura en un sistema ágil, elástico y de alta disponibilidad en la nube para reducir los gastos, aumentar la eficiencia y aprovechar las innovaciones. Para más información, consulte [Strategy for modernizing applications in the Nube de AWS](#).

evaluación de la preparación para la modernización

Evaluación que ayuda a determinar la preparación para la modernización de las aplicaciones de una organización; identifica los beneficios, los riesgos y las dependencias; y determina qué tan bien la organización puede soportar el estado futuro de esas aplicaciones. El resultado de la evaluación es un esquema de la arquitectura objetivo, una hoja de ruta que detalla las fases de desarrollo y los hitos del proceso de modernización y un plan de acción para abordar las brechas identificadas. Para más información, consulte [Evaluating modernization readiness for applications in the Nube de AWS](#).

aplicaciones monolíticas (monolitos)

Aplicaciones que se ejecutan como un único servicio con procesos estrechamente acoplados. Las aplicaciones monolíticas presentan varios inconvenientes. Si una característica de la aplicación experimenta un aumento en la demanda, se debe escalar toda la arquitectura. Agregar o mejorar las características de una aplicación monolítica también se vuelve más complejo a medida que crece la base de código. Para solucionar problemas con la aplicación, puede utilizar

una arquitectura de microservicios. Para obtener más información, consulte [Descomposición de monolitos en microservicios](#).

MPA

Consulte [Migration Portfolio Assessment](#).

MQTT

Consulte [Message Queuing Telemetry Transport](#).

clasificación multiclase

Un proceso que ayuda a generar predicciones para varias clases (predice uno de más de dos resultados). Por ejemplo, un modelo de ML podría preguntar “¿Este producto es un libro, un automóvil o un teléfono?” o “¿Qué categoría de productos es más interesante para este cliente?”.

infraestructura mutable

Modelo que actualiza y modifica la infraestructura actual para las cargas de trabajo de producción. Para mejorar la coherencia, la confiabilidad y la previsibilidad, el AWS Well-Architected Marco recomienda el uso de una [infraestructura inmutable](#) como práctica recomendada.

O

OAC

Consulte [control de acceso de origen](#).

OAI

Consulte [identidad de acceso de origen](#).

OCM

Consulte [administración del cambio organizacional](#).

migración fuera de línea

Método de migración en el que la carga de trabajo de origen se elimina durante el proceso de migración. Este método implica un tiempo de inactividad prolongado y, por lo general, se utiliza para cargas de trabajo pequeñas y no críticas.

OI

Consulte [integración de operaciones](#).

OLA

Consulte [acuerdo de nivel operativo](#).

migración en línea

Método de migración en el que la carga de trabajo de origen se copia al sistema de destino sin que se desconecte. Las aplicaciones que están conectadas a la carga de trabajo pueden seguir funcionando durante la migración. Este método implica un tiempo de inactividad nulo o mínimo y, por lo general, se utiliza para cargas de trabajo de producción críticas.

OPC-UA

Consulte [Open Process Communications: arquitectura unificada](#).

Comunicaciones de proceso abierto: arquitectura unificada () OPC-UA

Un protocolo de comunicación de máquina a máquina (M2M) para la automatización industrial. OPC-UA proporciona un estándar de interoperabilidad con esquemas de cifrado, autenticación y autorización de datos.

acuerdo de nivel operativo (OLA)

Acuerdo que aclara lo que los grupos de TI operativos se comprometen a ofrecerse entre sí, para respaldar un acuerdo de nivel de servicio (SLA).

revisión de la preparación operativa (ORR)

Lista de comprobación de preguntas y prácticas recomendadas asociadas que son útiles para comprender, evaluar, prevenir o reducir el alcance de los incidentes y posibles errores. Para obtener más información, consulte [las revisiones de preparación operativa \(ORR\)](#) en el AWS Well-Architected marco.

tecnología operativa (TO)

Sistemas de hardware y software que funcionan con el entorno físico para controlar las operaciones, los equipos y la infraestructura industriales. En el sector de la fabricación, la integración de los sistemas de TO y tecnología de la información (TI) es un enfoque clave para las transformaciones de la [industria 4.0](#).

integración de operaciones (OI)

Proceso de modernización de las operaciones en la nube, que implica la planificación de la preparación, la automatización y la integración. Para obtener más información, consulte la [Guía de integración de las operaciones](#).

registro de seguimiento organizativo

Un registro creado por y AWS CloudTrail que registra todos los eventos Cuentas de AWS de una organización AWS Organizations. Este registro de seguimiento se crea en cada Cuenta de AWS que forma parte de la organización y realiza un seguimiento de la actividad en cada cuenta. Para obtener más información, consulte [Crear un registro para una organización](#) en la CloudTrail documentación.

administración del cambio organizacional (OCM)

Marco para administrar las transformaciones empresariales importantes y disruptivas desde la perspectiva de las personas, la cultura y el liderazgo. La OCM ayuda a las empresas a prepararse para nuevos sistemas y estrategias y a realizar la transición a ellos, al acelerar la adopción de cambios, abordar los problemas de transición e impulsar cambios culturales y organizacionales. En la estrategia de AWS migración, este marco se denomina aceleración de personal, debido a la velocidad de cambio que requieren los proyectos de adopción de la nube. Para obtener más información, consulte la [Guía de OCM](#).

control de acceso de origen (OAC)

En CloudFront, una opción mejorada para restringir el acceso y proteger el contenido del Amazon Simple Storage Service (Amazon S3). El OAC admite todos los buckets de S3 Regiones de AWS, el cifrado del lado del servidor con AWS KMS (SSE-KMS) y DELETE las solicitudes PUT y dinámicas al bucket de S3.

identidad de acceso de origen (OAI)

En CloudFront, una opción para restringir el acceso y proteger el contenido de Amazon S3. Cuando utiliza OAI, CloudFront crea un principal con el que Amazon S3 puede autenticarse. Los directores autenticados solo pueden acceder al contenido de un bucket de S3 a través de una distribución específica. CloudFront Consulte también el [OAC](#), que proporciona un control de acceso más detallado y mejorado.

ORR

Consulte [revisión de la preparación operativa](#).

OT

Consulte [tecnología operativa](#).

VPC saliente (de salida)

En una arquitectura de AWS cuentas múltiples, una VPC que gestiona las conexiones de red que se inician desde una aplicación. La [Arquitectura de referencia de seguridad de AWS](#) recomienda

configurar su cuenta de red con VPC entrantes, salientes y de inspección para proteger la interfaz bidireccional entre su aplicación e Internet en general.

P

límite de permisos

Una política de administración de IAM que se adjunta a las entidades principales de IAM para establecer los permisos máximos que puede tener el usuario o el rol. Para obtener más información, consulte [Límites de permisos](#) en la documentación de IAM.

información de identificación personal (PII)

Información que, vista directamente o combinada con otros datos relacionados, puede utilizarse para deducir de manera razonable la identidad de una persona. Algunos ejemplos de información de identificación personal son los nombres, las direcciones y la información de contacto.

PII

Consulte [información de identificación personal](#).

manual de estrategias

Conjunto de pasos predefinidos que capturan el trabajo asociado a las migraciones, como la entrega de las funciones de operaciones principales en la nube. Un manual puede adoptar la forma de scripts, manuales de procedimientos automatizados o resúmenes de los procesos o pasos necesarios para operar un entorno modernizado.

PLC

Consulte [controlador lógico programable](#).

PLM

Consulte [administración del ciclo de vida del producto](#).

policy

Objeto que puede definir permisos (consulte [política basada en identidad](#)), especificar las condiciones de acceso (consulte [política basada en recursos](#)) o definir los permisos máximos para todas las cuentas de una organización de AWS Organizations (consulte [política de control de servicio](#)).

persistencia políglota

Elegir de forma independiente la tecnología de almacenamiento de datos de un microservicio en función de los patrones de acceso a los datos y otros requisitos. Si sus microservicios tienen la misma tecnología de almacenamiento de datos, pueden enfrentarse a desafíos de implementación o experimentar un rendimiento deficiente. Los microservicios se implementan más fácilmente y logran un mejor rendimiento y escalabilidad si utilizan el almacén de datos que mejor se adapte a sus necesidades.

evaluación de cartera

Proceso de detección, análisis y priorización de la cartera de aplicaciones para planificar la migración. Para obtener más información, consulte la [Evaluación de la preparación para la migración](#).

predicate

Condición de consulta que devuelve `true` o `false`. En general, se encuentra en una cláusula `WHERE`.

inserción de predicados

Técnica de optimización de consultas en bases de datos que filtra los datos de la consulta antes de transferirlos. Esta técnica reduce la cantidad de datos de la base de datos relacional que se tienen que recuperar y procesar. Además, mejora el rendimiento de las consultas.

control preventivo

Un control de seguridad diseñado para evitar que ocurra un evento. Estos controles son la primera línea de defensa para evitar el acceso no autorizado o los cambios no deseados en la red. Para obtener más información, consulte [Controles preventivos](#) en Implementación de controles de seguridad en AWS.

entidad principal

Una entidad AWS que puede realizar acciones y acceder a los recursos. Esta entidad suele ser un usuario raíz para un Cuenta de AWS rol de IAM o un usuario. Para obtener más información, consulte Entidad principal en [Términos y conceptos de roles](#) en la documentación de IAM.

Privacidad desde el diseño

Enfoque de ingeniería de sistemas que tiene en cuenta la privacidad durante todo el proceso de desarrollo.

zonas alojadas privadas

Contenedor que aloja información acerca de cómo desea que responda Amazon Route 53 a las consultas de DNS de un dominio y sus subdominios en una o varias VPC. Para obtener más información, consulte [Uso de zonas alojadas privadas](#) en la documentación de Route 53.

control proactivo

[Control de seguridad](#) que se diseñó para evitar la implementación de recursos que no cumplan con la normativa. Estos controles analizan los recursos antes de aprovisionarlos. Si el recurso no cumple con los requisitos del control, no se aprovisiona. Para obtener más información, consulte la [guía de referencia de controles](#) en la AWS Control Tower documentación y consulte [Controles proactivos](#) en Implementación de controles de seguridad en AWS.

administración del ciclo de vida del producto (PLM)

Administración de los datos y los procesos de un producto a lo largo de todo su ciclo de vida, desde el diseño, el desarrollo y el lanzamiento, pasando por el crecimiento y la madurez, hasta la reducción de su uso y su retirada.

entorno de producción

Consulte [entorno](#).

controlador lógico programable (PLC)

En el sector de la fabricación, computadora adaptable y altamente fiable que supervisa las máquinas y automatiza los procesos de fabricación.

encadenamiento de peticiones

Uso de la salida de una petición de [LLM](#) como entrada para la siguiente petición a fin de generar mejores respuestas. Esta técnica se utiliza para dividir una tarea compleja en tareas secundarias o para refinar o ampliar de forma iterativa una respuesta preliminar. Ayuda a mejorar la precisión y la relevancia de las respuestas de un modelo y permite obtener resultados más detallados y personalizados.

seudonimización

El proceso de reemplazar los identificadores personales de un conjunto de datos por valores de marcadores de posición. La seudonimización puede ayudar a proteger la privacidad personal. Los datos seudonimizados siguen considerándose datos personales.

publish/subscribe (pub/sub)

Patrón que permite establecer comunicaciones asíncronas entre microservicios para mejorar la escalabilidad y la capacidad de respuesta. Por ejemplo, en un [MES](#) basado en microservicios, un microservicio puede publicar mensajes de eventos en un canal al que se pueden suscribir otros microservicios. El sistema puede agregar nuevos microservicios sin cambiar el servicio de publicación.

Q

plan de consulta

Serie de pasos, como instrucciones, que se utilizan para acceder a los datos de un sistema de base de datos relacional SQL.

regresión del plan de consulta

El optimizador de servicios de la base de datos elige un plan menos óptimo que antes de un cambio determinado en el entorno de la base de datos. Los cambios en estadísticas, restricciones, configuración del entorno, enlaces de parámetros de consultas y actualizaciones del motor de base de datos PostgreSQL pueden provocar una regresión del plan.

R

Matriz RACI

Consulte [responsable, fiable, consultada e informada \(RACI\)](#).

RAG

Consulte [generación aumentada por recuperación](#).

ransomware

Software malicioso que se ha diseñado para bloquear el acceso a un sistema informático o a los datos hasta que se efectúe un pago.

Matriz RASCI

Consulte [responsable, fiable, consultada e informada \(RACI\)](#).

RCAC

Consulte [control de acceso por filas y columnas](#).

réplica de lectura

Una copia de una base de datos que se utiliza con fines de solo lectura. Puede enrutar las consultas a la réplica de lectura para reducir la carga en la base de datos principal.

rediseñar

Consulte [Las 7 R](#).

objetivo de punto de recuperación (RPO)

La cantidad de tiempo máximo aceptable desde el último punto de recuperación de datos. Esto determina qué se considera una pérdida de datos aceptable entre el último punto de recuperación y la interrupción del servicio.

objetivo de tiempo de recuperación (RTO)

La demora máxima aceptable entre la interrupción del servicio y el restablecimiento del servicio.

refactorizar

Consulte [Las 7 R](#).

Region

Conjunto de AWS recursos en un área geográfica. Cada uno Región de AWS está aislado e independiente de los demás para proporcionar tolerancia a las fallas, estabilidad y resiliencia. Para más información, consulte [Specify which Regiones de AWS your account can use](#).

regresión

Una técnica de ML que predice un valor numérico. Por ejemplo, para resolver el problema de “¿A qué precio se venderá esta casa?”, un modelo de ML podría utilizar un modelo de regresión lineal para predecir el precio de venta de una vivienda en función de datos conocidos sobre ella (por ejemplo, los metros cuadrados).

volver a alojar

Consulte [Las 7 R](#).

versión

En un proceso de implementación, el acto de promover cambios en un entorno de producción.

reubicar

Consulte [Las 7 R](#).

redefinir la plataforma

Consulte [Las 7 R](#).

recomprar

Consulte [Las 7 R](#).

resiliencia

Capacidad de una aplicación para resistir interrupciones o recuperarse de ellas. Al planificar la resiliencia en la Nube de AWS, la [alta disponibilidad](#) y la [recuperación ante desastres](#) son consideraciones comunes. Para más información, consulte [Resiliencia en la Nube de AWS](#).

política basada en recursos

Una política asociada a un recurso, como un bucket de Amazon S3, un punto de conexión o una clave de cifrado. Este tipo de política especifica a qué entidades principales se les permite el acceso, las acciones compatibles y cualquier otra condición que deba cumplirse.

matriz responsable, confiable, consultada e informada (RACI)

Una matriz que define las funciones y responsabilidades de todas las partes involucradas en las actividades de migración y las operaciones de la nube. El nombre de la matriz se deriva de los tipos de responsabilidad definidos en la matriz: responsable (R), contable (A), consultado (C) e informado (I). El tipo de soporte (S) es opcional. Si incluye el soporte, la matriz se denomina matriz RASCI y, si la excluye, se denomina matriz RACI.

control receptivo

Un control de seguridad que se ha diseñado para corregir los eventos adversos o las desviaciones con respecto a su base de seguridad. Para obtener más información, consulte [Controles receptivos](#) en Implementación de controles de seguridad en AWS.

retain

Consulte [Las 7 R](#).

retirar

Consulte [Las 7 R](#).

Generación aumentada de recuperación (RAG)

Tecnología de [IA generativa](#) mediante la que un [LLM](#) hace referencia a un origen de datos autorizado que se encuentra fuera de sus orígenes de datos de entrenamiento antes de generar una respuesta. Por ejemplo, un modelo de RAG podría hacer una búsqueda semántica en la base de conocimientos o en los datos personalizados de una organización. Para más información, consulte [¿Qué es RAG \(generación aumentada por recuperación\)?](#)

rotación

Proceso mediante el que periódicamente se actualiza un [secreto](#) para que resulte más difícil que un atacante pueda acceder a las credenciales.

control de acceso por filas y columnas (RCAC)

El uso de expresiones SQL básicas y flexibles que tienen reglas de acceso definidas. El RCAC consta de permisos de fila y máscaras de columnas.

RPO

Consulte [objetivo de punto de recuperación](#).

RTO

Consulte [objetivo de tiempo de recuperación](#).

manual de procedimientos

Conjunto de procedimientos manuales o automatizados necesarios para realizar una tarea específica. Por lo general, se diseñan para agilizar las operaciones o los procedimientos repetitivos con altas tasas de error.

S

SAML 2.0

Un estándar abierto que utilizan muchos proveedores de identidad (IdPs). Esta función permite el inicio de sesión único (SSO) federado, de modo que los usuarios pueden iniciar sesión en la Consola de administración de AWS o llamar a las operaciones de la AWS API sin tener que crear un usuario en IAM para todos los miembros de la organización. Para obtener más información sobre la federación basada en SAML 2.0, consulte [Acerca de la federación basada en SAML 2.0](#) en la documentación de IAM.

SCADA

Consulte [control de supervisión y adquisición de datos](#).

SCP

Consulte [política de control de servicio](#).

secreta

En AWS Secrets Manager, información confidencial o restringida, como una contraseña o credenciales de usuario, que se almacena de forma cifrada. Se compone del valor del secreto y de sus metadatos. El valor del secreto puede ser binario, una sola cadena o varias cadenas. Para más información, consulte [What's in a Secrets Manager secret?](#) en la documentación de Secrets Manager.

seguridad desde el diseño

Enfoque de ingeniería de sistemas que tiene en cuenta la seguridad durante todo el proceso de desarrollo.

control de seguridad

Barrera de protección técnica o administrativa que impide, detecta o reduce la capacidad de un agente de amenazas para aprovechar una vulnerabilidad de seguridad. Existen cuatro tipos de controles de seguridad principales: [preventivos](#), [de detección](#), [de respuesta](#) y [proactivos](#).

refuerzo de la seguridad

Proceso de reducir la superficie expuesta a ataques para hacerla más resistente a los ataques. Esto puede incluir acciones, como la eliminación de los recursos que ya no se necesitan, la implementación de prácticas recomendadas de seguridad consistente en conceder privilegios mínimos o la desactivación de características innecesarias en los archivos de configuración.

sistema de información sobre seguridad y administración de eventos (SIEM)

Herramientas y servicios que combinan sistemas de administración de información sobre seguridad (SIM) y de administración de eventos de seguridad (SEM). Un sistema de SIEM recopila, monitorea y analiza los datos de servidores, redes, dispositivos y otras fuentes para detectar amenazas y brechas de seguridad y generar alertas.

automatización de la respuesta de seguridad

Acción predefinida y programada que está diseñada para responder automáticamente a un evento de seguridad o corregirlo. Estas automatizaciones sirven como controles de seguridad

[preventivos o adaptables](#) que le ayudan a implementar las mejores prácticas AWS de seguridad. La modificación de un grupo de seguridad de VPC, la aplicación de revisiones a una instancia de Amazon EC2 o la rotación de credenciales son algunos ejemplos de acciones de respuesta automatizadas.

cifrado del servidor

Cifrado de los datos en su destino, por parte de Servicio de AWS quien los recibe.

política de control de servicio (SCP)

Una política que proporciona un control centralizado de los permisos de todas las cuentas de una organización en AWS Organizations. Las SCP definen barreras de protección o establecen límites a las acciones que un administrador puede delegar en los usuarios o roles. Puede utilizar las SCP como listas de permitidos o rechazados, para especificar qué servicios o acciones se encuentra permitidos o prohibidos. Para obtener más información, consulte [las políticas de control del servicio](#) en la AWS Organizations documentación.

punto de enlace de servicio

La URL del punto de entrada de un Servicio de AWS. Para conectarse mediante programación a un servicio de destino, puede utilizar un punto de conexión. Para obtener más información, consulte [Puntos de conexión de Servicio de AWS](#) en Referencia general de AWS.

acuerdo de nivel de servicio (SLA)

Acuerdo que aclara lo que un equipo de TI se compromete a ofrecer a los clientes, como el tiempo de actividad y el rendimiento del servicio.

indicador de nivel de servicio (SLI)

Medición de un aspecto del rendimiento de un servicio, como la tasa de errores, la disponibilidad o el rendimiento.

objetivo de nivel de servicio (SLO)

Métrica objetivo que representa el estado de un servicio medido mediante un [indicador de nivel de servicio](#).

modelo de responsabilidad compartida

Un modelo que describe la responsabilidad con AWS la que compartes la seguridad y el cumplimiento de la nube. AWS es responsable de la seguridad de la nube, mientras que usted es responsable de la seguridad en la nube. Para obtener más información, consulte el [Modelo de responsabilidad compartida](#).

Shadow AI

Aplicaciones de [IA](#) no autorizadas creadas o utilizadas fuera de los canales regulados dentro de una organización.

SIEM

Consulte [sistema de administración de eventos e información de seguridad](#).

único punto de error (SPOF)

Error en un único componente crítico de una aplicación que puede interrumpir el sistema.

SLA

Consulte [acuerdo de nivel de servicio](#).

SLI

Consulte [indicador de nivel de servicio](#).

SLO

Consulte [objetivo de nivel de servicio](#).

modelo de dividir y sembrar

Un patrón para escalar y acelerar los proyectos de modernización. A medida que se definen las nuevas funciones y los lanzamientos de los productos, el equipo principal se divide para crear nuevos equipos de productos. Esto ayuda a ampliar las capacidades y los servicios de su organización, mejora la productividad de los desarrolladores y apoya la innovación rápida. Para más información, consulte [Phased approach to modernizing applications in the Nube de AWS](#).

SPOF

Consulte [único punto de error](#).

esquema en estrella

Estructura organizativa de una base de datos que utiliza una tabla de hechos de gran tamaño para almacenar datos transaccionales o medidos y una o varias tablas dimensionales más pequeñas para almacenar los atributos de los datos. Esta estructura está diseñada para utilizarse en un [almacén de datos](#) o con fines de inteligencia empresarial.

patrón de higo estrangulador

Un enfoque para modernizar los sistemas monolíticos mediante la reescritura y el reemplazo gradual de las funciones del sistema hasta que se pueda dismantelar el sistema heredado.

Este patrón utiliza la analogía de una higuera que crece hasta convertirse en un árbol estable y, finalmente, se apodera y reemplaza a su host. El patrón fue [presentado por Martin Fowler](#) como una forma de gestionar el riesgo al reescribir sistemas monolíticos. Para ver un ejemplo de cómo aplicar este patrón, consulte [Modernización gradual de los servicios web antiguos de Microsoft ASP.NET \(ASMX\) mediante contenedores y Amazon API Gateway](#).

subred

Un intervalo de direcciones IP en la VPC. Una subred debe residir en una sola zona de disponibilidad.

control de supervisión y adquisición de datos (SCADA)

En el sector de la fabricación, sistema que utiliza hardware y software para supervisar los activos físicos y las operaciones de producción.

cifrado simétrico

Un algoritmo de cifrado que utiliza la misma clave para cifrar y descifrar los datos.

pruebas sintéticas

Prueba de un sistema de manera que simule las interacciones de los usuarios para detectar posibles problemas o supervisar el rendimiento. Puede usar [Amazon CloudWatch Synthetics](#) para crear estas pruebas.

petición del sistema

Técnica para proporcionar contexto, instrucciones o pautas a un [LLM](#) para dirigir su comportamiento. Las peticiones del sistema ayudan a establecer el contexto y las reglas para las interacciones con los usuarios.

T

etiquetas

Key-value pares que actúan como metadatos para organizar sus AWS recursos. Las etiquetas pueden ayudar a administrar, identificar, organizar, buscar y filtrar recursos de . Para obtener más información, consulte [Etiquetado de los recursos de AWS](#).

variable de destino

El valor que intenta predecir en el ML supervisado. Esto también se conoce como variable de resultado. Por ejemplo, en un entorno de fabricación, la variable objetivo podría ser un defecto del producto.

lista de tareas

Herramienta que se utiliza para hacer un seguimiento del progreso mediante un manual de procedimientos. La lista de tareas contiene una descripción general del manual de procedimientos y una lista de las tareas generales que deben completarse. Para cada tarea general, se incluye la cantidad estimada de tiempo necesario, el propietario y el progreso.

entorno de prueba

Consulte [entorno](#).

entrenamiento

Proporcionar datos de los que pueda aprender su modelo de ML. Los datos de entrenamiento deben contener la respuesta correcta. El algoritmo de aprendizaje encuentra patrones en los datos de entrenamiento que asignan los atributos de los datos de entrada al destino (la respuesta que desea predecir). Genera un modelo de ML que captura estos patrones. Luego, el modelo de ML se puede utilizar para obtener predicciones sobre datos nuevos para los que no se conoce el destino.

herramienta

Una función o API que un [agente](#) puede invocar para realizar operaciones en sistemas externos.

puerta de enlace de tránsito

Centro de tránsito de red que puede utilizar para interconectar las VPC y las redes en las instalaciones. Para obtener más información, consulte [Qué es una pasarela de tránsito](#) en la AWS Transit Gateway documentación.

flujo de trabajo basado en enlaces troncales

Un enfoque en el que los desarrolladores crean y prueban características de forma local en una rama de característica y, a continuación, combinan esos cambios en la rama principal. Luego, la rama principal se adapta a los entornos de desarrollo, preproducción y producción, de forma secuencial.

acceso de confianza

Otorgar permisos a un servicio que especifique para realizar tareas en su organización AWS Organizations y en sus cuentas en su nombre. El servicio de confianza crea un rol vinculado al servicio en cada cuenta, cuando ese rol es necesario, para realizar las tareas de administración por usted. Para obtener más información, consulte [AWS Organizations Utilización con otros AWS servicios](#) en la AWS Organizations documentación.

ajuste

Cambiar aspectos de su proceso de formación a fin de mejorar la precisión del modelo de ML. Por ejemplo, puede entrenar el modelo de ML al generar un conjunto de etiquetas, incorporar etiquetas y, luego, repetir estos pasos varias veces con diferentes ajustes para optimizar el modelo.

equipo de dos pizzas

Un DevOps equipo pequeño al que puedes alimentar con dos pizzas. Un equipo formado por dos integrantes garantiza la mejor oportunidad posible de colaboración en el desarrollo de software.

U

incertidumbre

Un concepto que hace referencia a información imprecisa, incompleta o desconocida que puede socavar la fiabilidad de los modelos predictivos de ML. Hay dos tipos de incertidumbre: la incertidumbre epistémica se debe a datos limitados e incompletos, mientras que la incertidumbre aleatoria se debe al ruido y la aleatoriedad inherentes a los datos.

tareas indiferenciadas

También conocido como tareas arduas, es el trabajo que es necesario para crear y operar una aplicación, pero que no proporciona un valor directo al usuario final ni proporciona una ventaja competitiva. Algunos ejemplos de tareas indiferenciadas son la adquisición, el mantenimiento y la planificación de la capacidad.

entornos superiores

Consulte [entorno](#).

V

succión

Una operación de mantenimiento de bases de datos que implica limpiar después de las actualizaciones incrementales para recuperar espacio de almacenamiento y mejorar el rendimiento.

control de versión

Procesos y herramientas que realizan un seguimiento de los cambios, como los cambios en el código fuente de un repositorio.

Emparejamiento de VPC

Conexión entre dos VPC que permite enrutar el tráfico mediante direcciones IP privadas. Para obtener más información, consulte [¿Qué es una interconexión de VPC?](#) en la documentación de Amazon VPC.

vulnerabilidad

Defecto de software o hardware que pone en peligro la seguridad del sistema.

W

caché caliente

Un búfer caché que contiene datos actuales y relevantes a los que se accede con frecuencia. La instancia de base de datos puede leer desde la caché del búfer, lo que es más rápido que leer desde la memoria principal o el disco.

datos templados

Datos a los que el acceso es infrecuente. Al consultar este tipo de datos, normalmente se aceptan consultas moderadamente lentas.

función de ventana

Función SQL que hace un cálculo en un grupo de filas que se relacionan de alguna manera con el registro actual. Las funciones de ventana son útiles para las tareas de procesamiento, como calcular una media móvil o acceder al valor de las filas en función de la posición relativa de la fila actual.

carga de trabajo

Conjunto de recursos y código que ofrece valor comercial, como una aplicación orientada al cliente o un proceso de backend.

flujo de trabajo

Grupos funcionales de un proyecto de migración que son responsables de un conjunto específico de tareas. Cada flujo de trabajo es independiente, pero respalda a los demás flujos de trabajo del proyecto. Por ejemplo, el flujo de trabajo de la cartera es responsable de priorizar las aplicaciones, planificar las oleadas y recopilar los metadatos de migración. El flujo de trabajo de la cartera entrega estos recursos al flujo de trabajo de migración, que luego migra los servidores y las aplicaciones.

WORM

Consulte [escritura única y lectura múltiple](#).

WQF

Consulte [AWS Workload Qualification Framework](#).

escritura única y lectura múltiple (WORM)

Modelo de almacenamiento que escribe los datos una sola vez y evita que se eliminen o modifiquen. Los usuarios autorizados pueden leer los datos tantas veces como sea necesario, pero no los pueden cambiar. Esta infraestructura de almacenamiento de datos se considera [inmutable](#).

Z

ataque de día cero

Ataque, normalmente de malware, que se aprovecha de una [vulnerabilidad de día cero](#).

vulnerabilidad de día cero

Un defecto o una vulnerabilidad sin mitigación en un sistema de producción. Los agentes de amenazas pueden usar este tipo de vulnerabilidad para atacar el sistema. Los desarrolladores suelen darse cuenta de la vulnerabilidad a raíz del ataque.

peticiones desde cero

Proporcionar a un [LLM](#) instrucciones para llevar a cabo una tarea, pero sin ejemplos (pasos) que puedan ayudar a guiarlo. El LLM debe usar los conocimientos del entrenamiento previo para

llevar a cabo la tarea. La eficacia de la petición desde cero depende de la complejidad de la tarea y de la calidad de la petición. Consulte también [peticiones con pocos pasos](#).

aplicación zombi

Aplicación que utiliza un promedio de CPU y memoria menor al 5 por ciento. En un proyecto de migración, es habitual retirar estas aplicaciones.

Las traducciones son generadas a través de traducción automática. En caso de conflicto entre la traducción y la versión original de inglés, prevalecerá la versión en inglés.