



Economía de la IA agencial en AWS

AWS Guía prescriptiva



AWS Guía prescriptiva: Economía de la IA agencial 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
Destinatarios previstos	1
Objetivos	2
Acerca de esta serie de contenido	2
Comprender la economía de la IA de los agentes	3
Evaluación de tareas	3
Evaluación del impacto del riesgo	4
Retorno de la inversión	6
Medir el éxito y el ROI	7
Usa tu base	7
Establezca objetivos	7
Rastrea las métricas	7
Uso AgentOps	8
Evaluación de los costos de los procesos humanos	8
Costes laborales	9
Costes de rendimiento	9
Costos de tecnología	10
Costes de oportunidad	11
Costes por riesgos y defectos	11
Implementación de sistemas de inteligencia artificial de agencias	12
Incorporar la retroalimentación humana	13
Aprendizaje conductual	14
Aprendizaje continuo	14
Colaboración entre humanos e inteligencia artificial	15
Precios basados en los resultados	16
Modelo tradicional e inicial	16
Modelo basado en resultados	16
Usando AWS Marketplace	17
Caso práctico: operaciones de contratación	19
Escenario A	19
Estructura de costos base	20
Métricas operativas	21
Análisis de costos basado en el volumen	22
Análisis del ROI	22

Comparación de costos acumulados	22
Beneficios adicionales	23
Escenario B	23
Estructura de costes base	24
Métricas operativas	25
Análisis de costos basado en el volumen	25
Análisis del ROI	26
Comparación de costos acumulados	26
Comparación de escenarios	27
Conclusión y recursos	29
Recursos	29
Historial de documentos	31
Glosario	32
#	32
A	33
B	36
C	38
D	41
E	46
F	48
G	50
H	51
I	52
L	55
M	56
O	60
P	63
Q	66
R	66
S	69
T	73
U	75
V	76
W	76
Z	77
.....	Ixxix

Economía de la IA agencial en AWS

Hans Schabert y Prasanta Roy, Amazon Web Services

Enero de 2026 ([historial del documento](#))

Las organizaciones que adoptan sistemas de automatización y de inteligencia artificial agencial impulsados por la IA necesitan tomar decisiones económicas informadas entre el trabajo humano y los agentes inteligentes. Esto se vuelve fundamental para la sostenibilidad de las operaciones en la nube. Esta guía le ayuda a evaluar, implementar y optimizar las compensaciones económicas entre la fuerza laboral humana y los sistemas de IA de los agentes. AWS Puede maximizar el retorno de la inversión (ROI) y, al mismo tiempo, mantener la excelencia operativa.

Ningún sistema es 100% correcto. Este principio fundamental impulsa el análisis económico de los sistemas de IA humanos y agenciales. Las organizaciones deben ir más allá de las comparaciones simplistas de costos para evaluar el impacto económico total, los perfiles de riesgo, los requisitos de calidad de las decisiones y la creación de valor estratégico a largo plazo.

El comportamiento de los clientes está cambiando drásticamente, pasando de las tradicionales inversiones iniciales en tecnología a pay-per-outcome modelos que alinean los costos con los resultados empresariales. Esta transformación requiere nuevos enfoques para la evaluación, la implementación y la optimización de la colaboración entre personas y agentes.

El camino hacia el éxito sigue un patrón claro: comience con los trabajos adecuados, mida todo y escale lo que funcione. Las organizaciones que adoptan este enfoque logran una ventaja competitiva sostenible mediante la asignación inteligente de recursos y la automatización centrada en los resultados.

Destinatarios previstos

Esta guía está destinada a lo siguiente:

- Ejecutivos (CEOs CTOs,, CFOs) que toman decisiones de inversión estratégica
- Arquitectos empresariales que diseñan estrategias de automatización organizacional
- Profesionales de operaciones financieras que optimizan la gestión financiera en la nube
- Líderes tecnológicos que están evaluando los enfoques de implementación de la IA
- Líderes de unidades de negocio que desean entender el ROI de la automatización

- Profesionales de compras que están navegando por los nuevos modelos de precios de la IA

Para comprender los conceptos de esta guía, le recomendamos que consulte los [fundamentos de la IA de los agentes en. AWS](#)

Objetivos

Esta guía le ayuda a comprender lo siguiente:

- Cómo evaluar el potencial de automatización de los agentes en los trabajos
- Modelos económicos para comparar los costos de la mano de obra humana con las inversiones en sistemas de IA de los agentes
- Pay-per-outcome los modelos de precios y su impacto en la economía de los proyectos de IA
- Técnicas de medición para demostrar el ROI y gestionar el riesgo
- Estrategias de escalado que transforman los costos fijos en resultados variables

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

Comprender la economía de la IA de los agentes en AWS

Uno de los principios clave es determinar cuándo utilizar los agentes de IA y cuándo utilizar los métodos deterministas tradicionales. Las organizaciones deben evaluar sistemáticamente qué trabajos justifican la automatización de los agentes y cuáles deben utilizar la automatización tradicional o la operación humana continua. Esta decisión requiere comprender la relación entre las características de la tarea, la tolerancia al riesgo y el enfoque operativo.

Antes de decidirse a implementar una IA institucional, debe utilizar el marco de decisión para comprender el impacto económico. El marco de decisión incluye las tres preguntas clave siguientes:

1. [Evaluación de tareas](#): ¿es esta tarea adecuada para un agente de IA?
2. [Evaluación del impacto del riesgo](#): ¿cuáles son los riesgos involucrados?
3. [Retorno de la inversión](#): ¿será rentable?

Evaluación de tareas

Las tareas con reglas de decisión estandarizadas y de alta complejidad pueden beneficiarse de los enfoques de IA de los agentes. La automatización tradicional o la automatización robótica de procesos permiten realizar mejor las tareas sencillas y altamente estandarizadas. Los sistemas de IA de las agencias se destacan a la hora de razonar, comprender el contexto o tomar decisiones de forma adaptativa. Añaden valor más allá del procesamiento basado en reglas. Las implementaciones exitosas de la IA de las agencias requieren sistemas que sean capaces de aprender y adaptarse.

Tenga en cuenta los siguientes factores al evaluar una tarea:

- **Complejidad:** grado de razonamiento y comprensión del contexto requerido. Las tareas que requieren una comprensión del contexto, una interpretación matizada o respuestas adaptativas a las condiciones cambiantes favorecen los enfoques de los agentes frente a la automatización tradicional, mientras que las tareas puramente mecánicas o de cálculo pueden no requerir inteligencia de los agentes.
- **Estandarización:** presencia de patrones y reglas claros. Se recomienda la IA de la agencia si la tarea requiere una comprensión del contexto. Si no se necesita adaptación ni aprendizaje, considere la automatización tradicional.
- **Volumen:** frecuencia de la ejecución de las tareas. Se recomienda la IA agencial para las actividades autónomas. Se recomienda la automatización tradicional para tareas consistentes y

de gran volumen. Sin embargo, el volumen por sí solo no determina el enfoque. Las decisiones de bajo volumen y alto valor podrían justificar la asistencia de los agentes para mejorar la calidad de las decisiones en lugar de reducir los costos.

- Valor: impacto empresarial por cada tarea completada. Considere la posibilidad de utilizar la inteligencia artificial para obtener resultados de gran valor que requieren una capacidad autónoma similar a la humana. Considere la automatización tradicional para tareas repetitivas y consistentes, que se pueden realizar de manera determinista.

Evaluación del impacto del riesgo

En la actualidad, existen cuatro enfoques de despliegue de la IA por agentes: totalmente autónomo, controlado por humanos, copiloto o dirigido por humanos con el apoyo de un agente. Cada uno tiene su propio perfil de riesgo y tolerancia al error, y todos implican a personas de alguna manera. En la siguiente tabla se describen los detalles de riesgo de estos enfoques.

Nivel de autonomía	Perfil de riesgo	Tolerancia a errores	Ejemplos de casos de uso	Participación humana
Totalmente autónomo	Bajo riesgo	1-2% aceptable	<ul style="list-style-type: none"> • Categorización básica de datos • Enrutamiento de documentos • Generación de informes estándar 	<ul style="list-style-type: none"> • Supervisión mínima • Auditorías periódicas
Humano al tanto	Riesgo medio	Por debajo del 0,5%	<ul style="list-style-type: none"> • Redacte las respuestas • Moderación de contenido 	<ul style="list-style-type: none"> • Revisión periódica • Tratamiento de excepciones

			<ul style="list-style-type: none"> • Tramitación inicial de las reclamaciones 	<ul style="list-style-type: none"> • Garantía de calidad
Copiloto	Alto riesgo	Casi cero	<ul style="list-style-type: none"> • Insumos para la planificación estratégica • Evaluaciones de riesgos • Decisiones de inversión 	<ul style="list-style-type: none"> • El ser humano toma las decisiones finales • El agente proporciona recomendaciones
Dirigido por humanos con el apoyo de agentes	Riesgo crítico	Tolerancia cero	<ul style="list-style-type: none"> • Decisiones legales • Diagnóstico médico • Conformidad normativa 	<ul style="list-style-type: none"> • El ser humano impulsa el proceso • El agente proporciona únicamente investigaciones o análisis e información de apoyo

En la siguiente tabla se describen las consideraciones clave a la hora de elegir entre estos enfoques.

Consideración	Totalmente autónomo	¡Humano al tanto	¿Copiloto	Dirigido por humanos
Rentabilidad	Máxima	Alto	Medio	Bajo
Escalabilidad	Sin límite	Alto	Medio	Limitado

Velocidad de procesamiento	El más rápido	Fast	Medio	Lento
Gestión de riesgos	Basic	Mejorado	Fuerte	Más fuerte
Manejo de la complejidad	Tareas sencillas	Tareas moderadamente complejas	Tareas complejas	Tareas críticas

Este marco de consideración ayuda a las organizaciones a adaptar los niveles de autonomía a los perfiles de riesgo, escalar las operaciones de manera adecuada, equilibrar la eficiencia con el control, implementar una gobernanza adecuada y optimizar la asignación de recursos.

Retorno de la inversión

El cálculo del retorno de la inversión de los sistemas de IA de los agentes comienza con un análisis de costes exhaustivo. Las organizaciones deben calcular primero sus costes humanos actuales, incluidos el salario, las prestaciones y los gastos de espacio de trabajo, junto con los gastos específicos del proceso y los costes ocultos, como la formación, la cobertura y el tiempo de inactividad.

Para analizar el punto de equilibrio, las organizaciones deben tener en cuenta los costos de implementación, los gastos operativos continuos y el volumen necesario para justificar la inversión. También es importante tener en cuenta las variaciones estacionales y los beneficios de la curva de aprendizaje que surgen a medida que los sistemas maduran y mejoran con el tiempo.

Al evaluar a los agentes de IA, las organizaciones deben recordar que estos sistemas suelen tener costes iniciales más altos, pero costes por transacción más bajos en comparación con las operaciones humanas. Además, los agentes de IA han demostrado mejorar el rendimiento a lo largo del tiempo y ofrecen una mejor escalabilidad que los equipos humanos. Esto los hace cada vez más rentables a medida que se amplía el despliegue y se acumula experiencia operativa.

Medición del éxito y el ROI de los sistemas de IA de las agencias

Medir el éxito en la implementación de un sistema de IA de agencias requiere un enfoque sistemático. Esta sección proporciona una metodología clara para la evaluación y la optimización continua que utiliza los análisis existentes en lugar de empezar desde cero.

Paso 1: Usa tu base existente

Comience con una evaluación integral de los costos de acuerdo con las recomendaciones de la sección [Evaluación de los costos actuales del proceso](#). Esto proporciona una base operativa para sus cálculos de ROI. Como se describe en la sección de [evaluación del impacto del riesgo](#), elija entre los cuatro niveles de autonomía (totalmente autónoma, integrada por una persona, con enfoque de copiloto, dirigida por una persona y con el apoyo de un agente) para determinar los criterios de medición y los umbrales de tolerancia a errores adecuados para cada proceso.

Paso 2: Establecer objetivos de éxito claros

Establezca objetivos de arquitectura y éxito que hagan hincapié en los sistemas con capacidad de aprendizaje, tal y como se describe en la sección [Patrones de éxito para la implementación de sistemas de IA institucionales](#). Céntrese en la mejora continua en lugar de en el rendimiento estático. Establezca plazos de retorno de la inversión utilizando la metodología de análisis de rentabilidad demostrada en el [estudio de caso: comparación de los costes de la IA humana y de los agentes en las operaciones de contratación](#). Incluya puntos de decisión claros para despedir a los agentes que no rinden sus servicios.

Paso 3: Realice un seguimiento de las métricas clave

Supervise el desempeño financiero con respecto a su base de referencia establecida y realice un seguimiento de los ahorros de costos y las mejoras de valor estratégico. Mida las métricas operativas, incluidas las tasas de error dentro de los umbrales aceptables para el nivel de autonomía que elija, las mejoras en la velocidad de procesamiento y las ganancias de coherencia. Céntrese en los indicadores estratégicos que demuestren la capacidad de aprendizaje y la adaptación a lo largo del tiempo.

Paso 4: Usa AgentOps

Aplique el marco de aprendizaje continuo de la sección [Incorporar la retroalimentación humana a los sistemas de IA de las agencias](#) para optimizar la toma de decisiones mediante la integración sistemática de la retroalimentación humana. Cree sistemas de aprendizaje en tiempo real que incorporen los conocimientos humanos para mejorar el rendimiento. Supervise la transformación hacia modelos de negocio basados en los resultados, tal como se describe en [Transformación económica hacia una política de precios basada en los resultados para los sistemas de inteligencia artificial de las agencias](#). AWS

Evaluación de los costes actuales de los procesos humanos

Comprender los costos reales de los procesos es fundamental para tomar decisiones informadas sobre las inversiones en sistemas de IA de los agentes. En primer lugar, debe establecer una base de referencia precisa del coste actual de sus procesos, incluidos todos los gastos ocultos, las tasas de fallos y los costes de oportunidad. Esto le ayuda a desarrollar cálculos precisos del ROI y a tomar decisiones estratégicas. Esta evaluación integral de costos constituye la base fundamental para evaluar si los sistemas de IA de los agentes pueden ofrecer un valor real como compañeros productivos.

La evaluación de los costes de referencia es fundamental por las siguientes razones clave:

- **Precisión del ROI:** las bases de referencia de costos precisas respaldan las proyecciones realistas del ROI que tienen en cuenta todo el espectro de los gastos operativos actuales.
- **Estrategia de implementación de agentes:** la comprensión exhaustiva de los costos ayuda a las organizaciones a identificar los procesos más prometedores para el despliegue inicial del sistema de inteligencia artificial de los agentes.
- **Medición del rendimiento:** las bases de referencia establecidas proporcionan el marco de medición para hacer un seguimiento de los beneficios reales y proyectados de las implementaciones de la IA de los organismos.

Las organizaciones deben identificar y evaluar sistemáticamente todos los factores de costo que influyen en la economía de los procesos antes de comparar las alternativas humanas y de los agentes. Esta evaluación garantiza cálculos de referencia precisos al tener en cuenta los factores de costo obvios y ocultos. Hace especial hincapié en los costes de los fallos, las tasas históricas de fallos y las oportunidades de negocio perdidas, que representan el coste total real de los procesos actuales.

En esta sección se describe cómo recopilar datos en cada categoría de costos para establecer mediciones de referencia precisas para sus procesos actuales. Analiza las fuentes de información y proporciona ejemplos de las siguientes categorías de costos:

- [Costes laborales](#)
- [Costes de rendimiento humano y coherencia](#)
- [Costes de tecnología e infraestructura](#)
- [Costes de oportunidades de negocio perdidas](#)
- [Costes por riesgos y defectos](#)

Costes laborales

Extraiga 24 meses de datos de nómina que incluyan el salario base, las horas extras, los beneficios y los costos de capacitación. Utilice su sistema de información de recursos humanos (HRIS) para realizar un seguimiento de los gastos de contratación y las tasas de rotación. Los sistemas de seguimiento del tiempo revelan la productividad real en comparación con las horas programadas. Las plataformas de gestión del rendimiento muestran la correlación entre los niveles de cualificación y los costes de compensación. Calcule las tarifas horarias completas que se destinen a los gastos generales de administración.

El siguiente es un ejemplo de una lista de los factores que impulsan los costos de la mano de obra.

Impulsor de costes	Impacto empresarial
Compensación base	Entre 25 y 150\$ por hora con carga completa
Beneficios e impuestos sobre la nómina	Entre el 25 y el 40% del salario base
Formación y desarrollo	Del 5 al 15% del costo laboral anual
Gastos generales de administración	Del 15 al 25% del costo laboral directo

Costes de rendimiento humano y coherencia

Combine los datos de los sistemas de gestión de proyectos que muestran las variaciones en la finalización de las tareas con los sistemas de asistencia. Esto puede revelar los patrones de absentismo y los cambios estacionales. Las plataformas de servicio al cliente muestran los rangos

de rendimiento individuales mediante métricas de resolución, y los datos de gestión de las relaciones con los clientes (CRM) de ventas pueden mostrar las variaciones de eficiencia en el cierre de operaciones. Los sistemas de gestión de la calidad proporcionan índices de defectos y procesan los datos de conformidad de todos los equipos y ubicaciones. Los sistemas de flujo de trabajo registran los tiempos de finalización, los retrasos en la aprobación y la frecuencia con la que se tramitan las excepciones. Los análisis de la comunicación revelan la sobrecarga de coordinación a través de la frecuencia de las reuniones y los patrones de colaboración.

El siguiente es un ejemplo de una lista de los factores que influyen en los costes del rendimiento y la coherencia de las personas.

Impulsor de costes	Impacto empresarial
Fluctuaciones de productividad	Rango de rendimiento del 20 al 50%
Absentismo y cobertura	Se necesita entre un 15 y un 25% de capacidad adicional
Ciclos de fatiga y motivación	Variación de productividad del 10 al 30%
Inconsistencias en los procedimientos	Pérdida de eficiencia del 10 al 40%
Variaciones en el control de calidad	Entre el 10 y el 30% del coste total
Gastos generales de coordinación	Del 15 al 25% del costo operativo

Costes de tecnología e infraestructura

Las plataformas de administración de licencias muestran los costos y las tasas de utilización del software. El monitoreo de la infraestructura proporciona datos sobre el tiempo de actividad, las métricas de rendimiento y los costos de mantenimiento. Los sistemas de asistencia técnica rastrean los gastos generales de soporte y los problemas técnicos recurrentes. Los sistemas de administración de proveedores capturan los costos totales de las relaciones tecnológicas, incluidos los gastos de integración y el rendimiento a nivel de servicio.

El siguiente es un ejemplo de una lista de los factores de costo de la tecnología y la infraestructura.

Impulsor de costes	Impacto empresarial
--------------------	---------------------

Sistemas tecnológicos	Entre 50 y 500 dólares por usuario al mes
Espacio de trabajo y equipo	Entre 200 y 1000 dólares por empleado al mes

Costes de oportunidades de negocio perdidas

Las plataformas de CRM contienen documentación sobre los tiempos de respuesta de los clientes potenciales, las tasas de conversión y las oportunidades perdidas. La automatización del marketing muestra que el retraso en el seguimiento afecta a las conversiones de clientes potenciales. Los sistemas de atención al cliente revelan cómo los problemas operativos afectan a la satisfacción y la retención. El análisis de la competencia proporciona los requisitos de respuesta del mercado y datos sobre ganancias o pérdidas que relacionan el rendimiento operativo con los resultados de ingresos.

A continuación se muestra un ejemplo de una lista de los factores que provocan la pérdida de oportunidades de negocio.

Impulsor de costes	Impacto empresarial
Demoras en la respuesta del mercado	Ingresos por día de retraso
Restricciones de capacidad	Oportunidades de negocio perdidas
Asignación de recursos de innovación	Coste de oportunidad del trabajo rutinario
Demoras en la adquisición de clientes	Entre un 50 y un 90% de pérdida de plomo por una respuesta lenta

Costes por riesgos y defectos

La documentación de la póliza de seguro muestra los costos de las coberturas de responsabilidad general, responsabilidad profesional, compensación laboral y responsabilidad cibernética. Los informes de evaluación interna de riesgos identifican las vulnerabilidades operativas y los costos de mitigación asociados. Los sistemas de seguimiento de defectos documentan las fallas en los productos o servicios, incluidos los costos de detección, los gastos de reemplazo y las reclamaciones de garantía. Los cronogramas de reemplazo de activos muestran las tasas de fallas de los equipos y los costos de reemplazo. Los informes sobre incidentes de seguridad registran los accidentes laborales y las reclamaciones de indemnización laboral asociadas. Los planes de continuidad

empresarial detallan los costos del sistema de respaldo y las inversiones en recuperación ante desastres.

A continuación se muestra un ejemplo de una lista de los factores de costo relacionados con los riesgos y los defectos.

Impulsor de costes	Impacto empresarial
Costos de seguro	Del 1 al 5% del presupuesto operativo
Coste de los errores	Entre 50 000 y 5000 dólares por incidente de error
Impacto de un error humano	Entre el 2 y el 15% del costo operativo total
Tasas de error y reprocesamiento	Entre 1,5 y 4 veces el coste original de las correcciones

Patrones exitosos para implementar sistemas de IA de agencias en AWS

[El estado de la adopción de la IA empresarial](#) (informe ISG 2025) revela que el principal obstáculo para una implementación exitosa de la IA no es la capacidad técnica sino la brecha de aprendizaje. Este término se refiere a los sistemas que no pueden adaptarse, recordar el contexto o mejorar con el tiempo. Organismos que implementan herramientas de IA estáticas registran altas tasas de fracaso. Las siguientes son características comunes de los sistemas de IA de las agencias que logran el éxito:

- Memoria contextual: sistemas que conservan el historial de conversaciones y las preferencias del usuario
- Integración de comentarios: capacidad de aprender de las correcciones y mejorar el rendimiento
- Adaptación del flujo de trabajo: ajuste automático a los cambiantes requisitos empresariales
- Mejora continua: mejora medible a través de la experiencia operativa

Las organizaciones que logran implementaciones de IA exitosas suelen priorizar lo siguiente:

- Utilizar ecosistemas de socios integrales en lugar de desarrollar y explorar las capacidades de IA de forma independiente
- Sistemas con capacidad de aprendizaje en lugar de herramientas estáticas
- Centrarse en los resultados empresariales más que en la comparación de las características técnicas
- Integración del flujo de trabajo en lugar de herramientas independientes
- Adaptación continua en lugar de implementación puntual

[Estos patrones se alinean con muchas Servicio de AWS capacidades, en particular el acceso al modelo básico en Amazon Bedrock, la arquitectura basada en eventos y la supervisión integral que se AWS Lambda ofrece a través de Amazon. CloudWatch](#) Para obtener más información sobre cómo integrar la retroalimentación humana y los sistemas con capacidad de aprendizaje, consulte la sección [Cómo incorporar la retroalimentación humana en los sistemas de inteligencia artificial de los agentes de esta guía.](#)

Incorporar la retroalimentación humana en los sistemas de IA de las agencias

Ningún sistema es 100% exitoso y es inevitable que se produzca un fracaso. Cada fallo conlleva un coste de cambio. Ser humano al tanto es un enfoque de IA en el que la IA realiza una tarea, pero se requiere la intervención o aprobación humana. Este enfoque debe utilizarse cuando el coste del fallo es superior al coste de disponer de una human-in-the-loop solución.

El éxito de los sistemas de IA de los agentes depende fundamentalmente de la capacidad del agente para aprender y mejorar gracias a la retroalimentación humana. Se debe tener en cuenta el coste del esfuerzo humano, en función del nivel de esfuerzo requerido. A diferencia de las herramientas de automatización estáticas que ejecutan reglas predeterminadas, las human-in-the-loop soluciones cuentan con sistemas de agentes con capacidad de aprendizaje que crean una asociación dinámica entre los agentes autónomos y el ser humano. La experiencia humana mejora continuamente el rendimiento del agente, mientras que los agentes gestionan el procesamiento rutinario a escala. Este enfoque colaborativo transforma la implementación de la IA de un solo despliegue a un proceso de optimización continuo. El sistema se adapta a los patrones organizativos, internaliza los estándares de calidad y refina sus capacidades de toma de decisiones basándose en la experiencia operativa del mundo real. Al recopilar sistemáticamente las correcciones, las aprobaciones y los conocimientos

humanos, las organizaciones pueden crear agentes de IA que entiendan el contexto, reconozcan los patrones y se adapten cada vez más a los objetivos empresariales a lo largo del tiempo.

En el caso de las soluciones que no requieren la intervención o el apoyo humanos, no es necesario tener en cuenta los costes específicos de las personas en la economía de los agentes.

Aprendizaje conductual por parte de operadores humanos

Los operadores humanos proporcionan información fundamental que los sistemas de inteligencia artificial de las agencias pueden utilizar para aprender, adaptarse y mejorar sus respuestas con el tiempo. Este circuito de retroalimentación crea un entorno de colaboración en el que la experiencia humana mejora la capacidad de los agentes mientras los agentes se encargan del procesamiento rutinario.

Mediante el reconocimiento de patrones de comportamiento humano, los agentes aprenden de los patrones de interacción humana para reflejar los enfoques de comunicación exitosos. Esto les ayuda a adaptarse a los patrones de decisión de la organización y a los niveles de tolerancia al riesgo. Los sistemas internalizan las expectativas de calidad mediante correcciones y aprobaciones humanas. También pueden aprender las respuestas adecuadas para los diferentes segmentos de clientes y contextos empresariales.

Los mecanismos efectivos de recopilación de comentarios capturan sistemáticamente las modificaciones y modificaciones humanas en las respuestas de los agentes. Analizan lo que los revisores humanos aprueban, rechazan o modifican en las recomendaciones de los agentes. Al comprender por qué ciertos casos requieren la intervención humana e incorporar la evaluación humana del desempeño de los agentes en diferentes escenarios y niveles de complejidad, estos sistemas perfeccionan continuamente sus capacidades para alinearlas mejor con los estándares y expectativas de la organización.

Operaciones de aprendizaje continuo

La integración del aprendizaje en tiempo real permite a los sistemas de inteligencia artificial de los agentes incorporar la retroalimentación humana y mejorar las respuestas de los agentes de forma inmediata mediante la actualización dinámica del modelo. Estos sistemas utilizan los conocimientos humanos para identificar nuevos patrones y casos extremos. Esto mejora sus capacidades de reconocimiento de patrones y, al mismo tiempo, desarrolla la memoria organizacional a través de experiencias de aprendizaje guiadas por humanos. El perfeccionamiento continuo basado en los comentarios de las personas y los operadores y en los resultados empresariales impulsa la optimización continua del rendimiento.

La formación guiada por personas captura el conocimiento experto para mejorar las capacidades de toma de decisiones de los agentes. Transfiere la experiencia fundamental de los operadores experimentados al sistema de IA. Mediante el aprendizaje basado en escenarios, los sistemas utilizan ejemplos creados por humanos para mejorar su manejo de situaciones complejas. También alinean los estándares de desempeño de los agentes con las expectativas de calidad humana mediante la calibración de la calidad. Este enfoque incorpora los conocimientos humanos sobre la cultura organizacional y las expectativas de los clientes. Esta adaptación cultural ayuda a los agentes a responder adecuadamente en diferentes contextos.

Excelencia operativa con la colaboración entre humanos e inteligencia artificial

La optimización automatizada y consciente de los riesgos permite una evaluación continua de las condiciones de operación y la probabilidad de error con supervisión humana en situaciones de alto riesgo. Esto ayuda a los sistemas a aprender de las evaluaciones del riesgo humano y a mejorar la toma de decisiones en el futuro. [Amazon Bedrock](#) proporciona acceso a varios modelos de base con diferentes capacidades y perfiles de costes. Esto permite un enrutamiento inteligente que tiene en cuenta los perfiles de costo y riesgo y, al mismo tiempo, incorpora la retroalimentación humana para optimizar la selección de modelos. El ajuste del rendimiento equilibra la eficiencia con la minimización de la tasa de errores al incorporar la opinión humana sobre los estándares de calidad y las compensaciones de rendimiento aceptables. Las decisiones automatizadas tienen en cuenta el coste total de propiedad ajustado al riesgo. Los operadores brindan orientación sobre la tolerancia al riesgo organizacional y la ponderación de las prioridades empresariales. Esto le ayuda a optimizar los costos y, al mismo tiempo, a alinearse con los objetivos de la organización.

Los sistemas de aprendizaje mejorados por las personas priorizan las aportaciones humanas teniendo en cuenta el impacto de los errores y las consecuencias empresariales. Esto crea sistemas de aprendizaje que comprenden tanto la precisión técnica como el contexto empresarial mediante comentarios ponderados en función del riesgo. El análisis periódico del rendimiento incorpora métricas de riesgo y análisis del coste de los errores, y los conocimientos humanos proporcionan un contexto que los sistemas automatizados no pueden captar. El desarrollo de las mejores prácticas hace hincapié en la gestión de riesgos y la prevención de errores al combinar el reconocimiento automático de patrones con la experiencia y el juicio humanos. El desarrollo de capacidades organizativas mediante programas de formación desarrolla tanto las habilidades humanas para gestionar los sistemas de IA de los agentes como las capacidades de los agentes para apoyar la toma de decisiones humanas. Esto garantiza un enfoque integral de la colaboración entre el ser humano y la IA que fortalece ambos componentes de la asociación.

Transformación económica hacia una fijación de precios basada en los resultados para los sistemas de IA de las agencias en AWS

El cambio de los modelos tradicionales de costes fijos a los precios basados en los resultados representa una transformación fundamental en la forma en que las organizaciones estructuran sus operaciones económicas y gestionan el riesgo. Esta transformación sirve como vía para la modernización constante de los procesos existentes y, al mismo tiempo, financia la transformación de la IA institucional. Permite a las organizaciones pasar de operaciones estáticas que consumen muchos recursos a modelos de negocio dinámicos y basados en los resultados.

Modelo tradicional e inicial

Los departamentos suelen funcionar como centros de costos con costos laborales directos que se financian mediante la asignación de costos. Por lo general, las organizaciones desean reducir esta asignación de costos. Si el proceso no está modernizado, el departamento debe ofrecer los mismos resultados con una fuerza laboral más reducida. Esto suele degradar la calidad. Los modelos de negocio tradicionales plantean desafíos importantes, entre los que se incluyen:

- Escalamiento lineal de los costos con aumentos de volumen: esto requiere que las organizaciones contraten personal adicional para gestionar el aumento del volumen.
- Compromisos de costos fijos: persisten independientemente del rendimiento empresarial y la eficiencia de los procesos.
- Planificación avanzada: la flexibilidad limitada durante las recesiones económicas y las limitaciones de capacidad requiere una planificación previa.
- Ciclo de degradación de la calidad: la reducción de los presupuestos conduce a una disminución de la calidad del servicio cuando los costos se reducen sin mejoras en los procesos.

Modelo basado en resultados

Los modelos modernos basados en los resultados vinculan los pagos directamente a los resultados empresariales mensurables, como las contrataciones realizadas con éxito, las métricas de calidad alcanzadas, las mejoras en la eficiencia de los procesos o los aumentos de productividad logrados. Esto traslada fundamentalmente el riesgo financiero de las unidades de negocio a los proveedores de servicios, al tiempo que crea una alineación natural de los incentivos. Los principales beneficios de un modelo basado en los resultados son los siguientes:

- Los costos aumentan directamente con el valor empresarial generado
- Alineación natural entre los gastos operativos y los ingresos
- Flexibilidad para ajustar la capacidad en función de las condiciones del mercado
- Pay-per-success los modelos reducen el riesgo financiero al desplazar la exposición financiera de una inversión inicial a un rendimiento operativo continuo
- Céntrese en sistemas con capacidad de aprendizaje que mejoren con el tiempo, en lugar de en alternativas estáticas

Esta transformación va mucho más allá de los centros de costos internos y redefine de manera fundamental la forma en que las organizaciones interactúan con los socios y proveedores de servicios externos. Al aplicar precios basados en los resultados a las colaboraciones con los socios, las organizaciones pueden impulsar mejoras de calidad a largo plazo y reducir los costos, al tiempo que hacen hincapié indirectamente en la modernización de la IA de los agentes.

Las organizaciones pueden experimentar rápidamente, medir el rendimiento con claridad y escalar en función del valor empresarial real generado en lugar de los compromisos tradicionales de recursos fijos. Este enfoque permite lo siguiente:

- Evolución de la relación con los proveedores: los socios se centran en el éxito de los clientes y no solo en la prestación de servicios.
- Métricas de resultados estandarizadas: simplifique los procesos de adquisición en varios proveedores.
- Capacidad de respuesta del mercado: adáptese rápidamente a las condiciones cambiantes del mercado y a las necesidades de los clientes.
- Ventaja competitiva: utilización superior de los recursos y capacidades operativas mejoradas.
- Alianzas impulsadas por la calidad: la colaboración a largo plazo se centra en la mejora continua y en los resultados mensurables.

Utilización AWS Marketplace como facilitador pay-per-outcome

El factor clave de esta transformación es [AWS Marketplace](#) que sirve como vehículo de transacciones para el trabajo de los agentes y la fijación de precios basada en los resultados. Proporciona acceso a cientos de agentes y soluciones de agentes de IA prediseñados con modelos de precios transparentes y basados en el uso. Puede ayudar a eliminar los costos iniciales de licencias, reducir

la complejidad de la implementación y permitir que las organizaciones se centren en sistemas aptos para el aprendizaje que se adaptan y mejoran con el tiempo, en lugar de en alternativas estáticas

Su uso AWS Marketplace puede proporcionar los siguientes beneficios:

- Experimentación rápida: pruebe varias soluciones sin una inversión de capital significativa
- Precios transparentes: costes basados en el uso con una clara atribución a los resultados empresariales
- Soluciones probadas: acceso a agentes probados por proveedores con experiencia
- Integración integrada: conectividad perfecta con los sistemas existentes Servicios de AWS
- Mitigación de riesgos: posibilidad de cambiar de proveedor en función del rendimiento
- Acceso a la capacidad de aprendizaje: disponibilidad de sistemas adaptativos sin costos de desarrollo interno

Este enfoque permite a las organizaciones comparar múltiples opciones en función de la obtención de resultados y las capacidades de aprendizaje, en lugar de en las listas de características. También puede ayudarlo a establecer criterios de éxito y metodologías de medición claros y a negociar precios basados en los resultados que estén vinculados a los resultados empresariales y a la mejora del sistema. Al financiar la transformación de la IA de los organismos mediante modelos basados en los resultados, las organizaciones pueden modernizar sus procesos de forma continua y, al mismo tiempo, pagar únicamente por mejoras cuantificables y resultados satisfactorios.

Caso práctico: comparación de los costes de la IA humana y de los agentes en las operaciones de contratación

Las operaciones de contratación ofrecen un caso práctico convincente para evaluar las desventajas económicas entre los sistemas de IA humanos y los de los agentes, pero el cálculo del ROI depende en gran medida de la base operativa actual. Las organizaciones que evalúan las inversiones de los agentes en IA suelen plantearse una pregunta fundamental: «¿Y si simplemente optimizamos nuestros procesos humanos existentes?» Para abordar esto directamente, este análisis presenta dos escenarios distintos que abarcan el rango de la eficiencia operativa humana.

[El escenario A](#) modela los tiempos de revisión de un currículum vitae (CV) o de 45 minutos. [El escenario B](#) demuestra la optimización de las operaciones humanas en 15 minutos por aplicación, lo que supone una mejora de la eficiencia del 66%. Por ejemplo, esta mejora podría lograrse mediante procesos simplificados, reclutadores experimentados o herramientas especializadas.

Al comparar capacidades idénticas de un sistema de agentes con estas diferentes líneas de base de rendimiento humano, revelamos cómo la eficiencia de los procesos actuales afecta a los cálculos del ROI, a los plazos de equilibrio y a las decisiones de implementación estratégica. Este enfoque de dos escenarios sirve para múltiples propósitos. Impide que las organizaciones descarten la IA de los agentes al suponer que la optimización de los procesos por sí sola es suficiente. También ayuda a las organizaciones con procesos ya eficientes a entender su economía específica. Además, estos escenarios ponen de relieve cuándo las ventajas no financieras, como la disponibilidad ininterrumpida y la escalabilidad, se convierten en los principales factores de decisión. Comprender estas dinámicas económicas a través de diferentes parámetros de eficiencia permite a las organizaciones tomar decisiones informadas sobre dónde y cuándo implementar los sistemas de inteligencia artificial de los agentes para lograr el máximo impacto empresarial.

Escenario A: tiempo de proyección de 45 minutos

El escenario A representa las operaciones de contratación en las que los reclutadores humanos dedican 45 minutos a revisar cada currículum. Este escenario modela a un reclutador de nivel medio con un coste total anual de 112.250\$. Este reclutador procesa las solicitudes durante el horario laboral estándar con características de desempeño humano típicas. Por el contrario, el sistema de IA de los agentes requiere una inversión inicial de 23 000 dólares para el desarrollo, la personalización y la integración de ATS, y tiene unos costes operativos mensuales mínimos de 500 dólares para la infraestructura en la nube. El agente procesa las aplicaciones en solo 5 minutos y está disponible

las 24 horas del día, los 7 días de la semana, lo que le permite alcanzar una tasa de error del 2% y una capacidad mensual superior a las 8.600 aplicaciones. Se trata de una brecha de eficiencia considerable, ya que el agente opera 9 veces más rápido por aplicación y tiene una capacidad mensual 39 veces mayor. En esta sección se examina el análisis de la estructura de costes, las métricas operativas, las comparaciones basadas en el volumen y los cálculos del ROI acumulado durante los primeros seis meses de funcionamiento.

Estructura de costos base

La siguiente tabla muestra los costos de configuración iniciales para el escenario A.

Componente	Operaciones humanas	Sistema de IA agencial
Desarrollo y personalización de agentes	N/A	15.000 dólares
Integración del sistema de seguimiento de candidatos (ATS)	N/A	5.000\$
Capacitación y optimización	N/A	3.000\$
Coste total de configuración inicial	0\$	23.000 DÓLARES

La siguiente tabla muestra los costos fijos anuales para el escenario A.

Componente	Operaciones humanas	Sistema de IA agencial
Salario base	65.000\$	N/A
Prestaciones (30%)	19.500 DÓLARES	N/A
Espacio de trabajo y equipo	12.000\$	N/A
Supervisión de la gestión (15%)	9.750 DÓLARES	N/A
Capacitación y desarrollo	6.000 dólares	N/A

Coste fijo anual total	112.250\$	N/D
------------------------	-----------	-----

La siguiente tabla muestra los costos operativos mensuales para el escenario A.

Componente	Operaciones humanas	Sistema de IA agencial
Computación en la nube	N/A	\$200
Almacenamiento	N/A	100 USD
Operaciones de base de datos	N/A	100 USD
Supervisión	N/A	100 USD
Coste fijo mensual total	9.354\$	500 DÓLARES

Métricas operativas

La siguiente tabla muestra las métricas operativas del escenario A.

Métrica	Operaciones humanas	Sistema de IA agencial
Tiempo de procesamiento por aplicación	45 minutos	5 minutos
Capacidad horaria	1,33 aplicaciones	12 aplicaciones
Capacidad diaria (24 horas)	10-11 aplicaciones	288 aplicaciones
Capacidad mensual	220 aplicaciones	8.640 aplicaciones
Coste por aplicación	45\$	2,50 DÓLARES
Coste por contratación exitosa	2.200\$	\$125
Tasa de error	5%	2%
Coste de corrección de errores	90\$ por error	45\$ por escalada

Análisis de costos basado en el volumen

La siguiente tabla muestra un análisis de costes basado en el volumen para el escenario A. En este ejemplo, el coste del sistema de IA de la agencia incluye los costes fijos y los costes de configuración amortizados, que ascienden a 1.917\$ al mes durante 12 meses.

Volumen mensual	Coste humano	Coste del sistema de IA de la agencia	Ahorros mensuales
100 aplicaciones	4.500\$	750 DÓLARES	3.750 DÓLARES
500 solicitudes	22.500 DÓLARES	2.667 DÓLARES	19.833 DÓLARES
1000 solicitudes	45.000\$	4.917 DÓLARES	40.083 DÓLARES

Análisis del ROI

La siguiente tabla muestra un análisis del ROI para el escenario A que se basa en el procesamiento de 500 solicitudes al mes.

Métrica	Valor
Coste humano mensual	22.500 dólares
Coste mensual del agente	2.667 dólares
Ahorros mensuales	19.833\$
Ahorros anuales	237.996 DÓLARES
Período de equilibrio	1,16 meses

Comparación de costos acumulados

La siguiente tabla muestra una comparación de costos acumulados para el escenario A durante los primeros seis meses, suponiendo 500 solicitudes por mes.

Mes	Coste humano	Coste del sistema de IA de la agencia	Ahorros acumulados
1	22.500 dólares	25.667 DÓLARES	-3167 DÓLARES
2	45.000 DÓLARES	28.334 DÓLARES	16.666 DÓLARES
3	67.500 DÓLARES	31.001 DÓLARES	36.499 DÓLARES
4	90.000 DÓLARES	33.668 DÓLARES	56.332 DÓLARES
5	112.500 DÓLARES	36.335 DÓLARES	76.165 DÓLARES
6	135.000 DÓLARES	39.002 DÓLARES	95.998 DÓLARES

Ventajas adicionales del sistema de inteligencia artificial para agentes

Los siguientes son beneficios adicionales que ofrece el sistema de IA de la agencia en el escenario A:

- Escalabilidad: puede gestionar picos de volumen sin coste adicional
- Disponibilidad: funcionamiento ininterrumpido con respuesta inmediata
- Coherencia: aplicación uniforme de los criterios de selección
- Eficiencia de tiempo: reducida significativamente time-to-hire
- Experiencia de usuario: comentarios instantáneos a los candidatos

Escenario B: tiempo de proyección de 15 minutos

Los modelos del escenario B optimizaron las operaciones de contratación, mientras que los reclutadores humanos simplificaron su proceso de selección a 15 minutos por solicitud. Esto representa una mejora de la eficiencia del 66% con respecto al escenario A. Este escenario mantiene el mismo coste anual total de 112.250 dólares para un reclutador de nivel medio. Sin embargo, demuestra una mejora significativa de la productividad humana, ya que la capacidad diaria aumenta a 32 solicitudes en un turno de 8 horas y el rendimiento mensual alcanza las 660 solicitudes. La mejora de la eficiencia humana reduce el coste por aplicación de 45 a 15 dólares, lo que reduce la brecha económica con el sistema de IA de los agentes. Sin embargo, el agente mantiene sus

ventajas estructurales: un tiempo de procesamiento de 5 minutos, disponibilidad ininterrumpida para 288 aplicaciones diarias, una tasa de error inferior al 2% en comparación con el 5% humano y una capacidad mensual superior a las 8.600 aplicaciones. Si bien esta mejora de la eficiencia amplía el período de equilibrio de 1,16 meses a 4,76 meses y reduce los ahorros mensuales de 19 833 a 4 833 dólares, el análisis revela que los sistemas de agentes siguen siendo económicamente viables incluso cuando compiten con operaciones humanas altamente optimizadas, una información fundamental para las organizaciones que evalúan si sus niveles actuales de eficiencia de los procesos justifican la inversión en inteligencia artificial de los agentes.

Estructura de costes base

La siguiente tabla muestra los costos fijos anuales para el escenario B.

Componente	Operaciones humanas	Sistema de IA agencial
Salario base	65.000\$	N/A
Prestaciones (30%)	19.500 DÓLARES	N/A
Espacio de trabajo y equipo	12.000\$	N/A
Supervisión de la gestión (15%)	9.750 DÓLARES	N/A
Capacitación y desarrollo	6.000 dólares	N/A
Costo fijo anual total	112.250\$	N/D

La siguiente tabla muestra los costos de implementación para el escenario B.

Componente	Operaciones humanas	Sistema de IA agencial
Configuración inicial de	N/A	23.000\$
Costos fijos mensuales	9.354\$	500 DÓLARES

Métricas operativas

La siguiente tabla muestra las métricas operativas del escenario B.

Métrica	Operaciones humanas	Sistema de IA agencial
Tiempo de procesamiento por aplicación	15 minutos	5 minutos
Capacidad horaria	4 aplicaciones	12 aplicaciones
Capacidad diaria (turno de 8 horas)	32 aplicaciones	288 aplicaciones
Capacidad mensual	660 aplicaciones	8.640 aplicaciones
Coste por aplicación	15 dólares	2,50 DÓLARES
Coste por contratación exitosa	2.200\$	\$125
Tasa de error	5%	2%
Coste de corrección de errores	30\$ por error	45\$ por escalada

Análisis de costos basado en el volumen

La siguiente tabla muestra un análisis de costes basado en el volumen para el escenario B. En este ejemplo, el coste del sistema de IA de la agencia incluye los costes fijos y los costes de configuración amortizados de 1.917\$ al mes durante 12 meses.

Volumen mensual	Coste humano	Coste del sistema de IA de la agencia	Ahorros mensuales
100 aplicaciones	1.500\$	750 DÓLARES	750 DÓLARES
500 aplicaciones	7.500 DÓLARES	2.667 DÓLARES	4.833 DÓLARES
1000 solicitudes	15.000 dólares	4.917 DÓLARES	10.083 DÓLARES

Análisis del ROI

La siguiente tabla muestra un análisis del ROI para el escenario B que se basa en el procesamiento de 500 solicitudes al mes.

Métrica	Valor
Coste humano mensual	7.500 dólares
Coste mensual del sistema de IA de la agencia	2.667\$
Ahorros mensuales	4.833\$
Ahorros anuales	57.996 DÓLARES
Período de equilibrio	4,76 meses

Comparación de costos acumulados

La siguiente tabla muestra una comparación de costos acumulados para el escenario B durante los primeros seis meses, suponiendo 500 solicitudes por mes.

Mes	Coste humano	Coste del sistema de IA de la agencia	Ahorros acumulados
1	7.500\$	25.667 DÓLARES	- 18.167 DÓLARES
2	15.000 DÓLARES	28.334 DÓLARES	- 13.334 DÓLARES
3	22.500 DÓLARES	31.001 DÓLARES	-8.501 DÓLARES
4	30 000 DÓLARES	33.668 DÓLARES	-3668 DÓLARES
5	37.500 DÓLARES	36.335 DÓLARES	1.165 DÓLARES
6	45.000 DÓLARES	39.002 DÓLARES	5.998 DÓLARES

Comparación de costos y beneficios para cada escenario

Métrica	Escenario A	Escenario B	Impact
Tiempo de proyección	45 minutos	15 minutos	mejora del 66%
Capacidad diaria	De 10 a 11 aplicaciones	32 aplicaciones	Aumento del 200%
Coste por aplicación	45\$	15 dólares	Reducción del 66%
Ahorros mensuales (500 solicitudes)	19.833 DÓLARES	4.833 DÓLARES	Disminución del 76%
Período de equilibrio	1,16 meses	4,76 meses	Un 310% más

El escenario B demuestra un aumento significativo de la eficiencia en las operaciones humanas, con mejoras en el tiempo de procesamiento que aumentan la capacidad sin personal adicional y reducen sustancialmente el costo por aplicación. Sin embargo, el impacto financiero revela un panorama más matizado: si bien el ROI sigue siendo positivo, las organizaciones se enfrentan a un período de equilibrio prolongado y a una reducción de los ahorros mensuales en comparación con el escenario A. Estos resultados destacan los factores de decisión críticos para la implementación: el sistema de agentes sigue siendo financieramente viable incluso con operaciones humanas optimizadas, pero las organizaciones deben adoptar una perspectiva de inversión a más largo plazo y considerar cuidadosamente las fluctuaciones de volumen y las necesidades de escalabilidad al evaluar los plazos de implementación y los retornos esperados.

Sin embargo, el sistema de IA de la agencia aún mantiene ventajas operativas críticas que van más allá del simple ahorro de costos. Ofrece disponibilidad ininterrumpida para la participación inmediata de los candidatos, independientemente de las zonas horarias o el horario laboral. Ofrece una calidad de selección uniforme al aplicar criterios uniformes a cada solicitud y se adapta a los picos de volumen sin incurrir en costes adicionales. Ofrece una respuesta inmediata a los candidatos, lo que mejora la imagen de la empresa y la experiencia del candidato, y funciona con un factor de fatiga cero, lo que garantiza el mismo rendimiento de alta calidad en la primera solicitud que en la milésima.

Los errores humanos suelen ser el resultado de la fatiga, la distracción o la falta de conocimiento y, a menudo, implican una falta de comunicación o información incorrecta. Los errores del sistema de inteligencia artificial de las agencias suelen deberse a casos extremos, entradas ambiguas

o limitaciones de los datos de entrenamiento. Estos errores tienden a ser de naturaleza más consistente.

Las métricas de calidad y experiencia revelan un equilibrio claro entre las capacidades humanas y las de los agentes:

- Satisfacción del cliente: las personas destacan por la empatía y la resolución de problemas complejos, y los agentes proporcionan información coherente y precisa para las consultas rutinarias.
- Tiempo de respuesta: el tiempo de respuesta favorece a los agentes con una disponibilidad inmediata las 24 horas del día, los 7 días. Los empleados prestan asistencia en horario laboral ante posibles retrasos en las colas.
- Coherencia: los agentes ofrecen respuestas idénticas a consultas similares. Los seres humanos pueden variar en cuanto al enfoque y la aplicación de los conocimientos.
- Manejo de la escalada: los problemas complejos que requieren juicio, creatividad o inteligencia emocional siguen siendo puntos fuertes de los seres humanos.

Conclusión y recursos

La economía de los sistemas humanos en comparación con los sistemas de IA de los agentes representa más que una decisión tecnológica. Refleja una transformación fundamental en la forma en que las organizaciones crean valor, gestionan el riesgo y logran una ventaja competitiva. El éxito requiere una evaluación sistemática de las características del puesto, una medición exhaustiva de los resultados (incluidos los factores de riesgo) y un escalamiento estratégico basado en resultados comprobados.

[El estado de la adopción de la IA empresarial](#) (informe ISG 2025) revela que la mayoría de las implementaciones de IA fracasan debido a brechas de aprendizaje, es decir, sistemas que no pueden adaptarse, recordar el contexto o mejorar con el tiempo. Las organizaciones que logran el éxito se centran en sistemas con capacidad de aprendizaje que se integran profundamente en los flujos de trabajo y demuestran una mejora continua a través de la retroalimentación humana y la experiencia operativa.

Las organizaciones que entiendan estos principios (empezando por los puestos adecuados, dividiendo los trabajos en tareas, midiéndolo todo, incluido el impacto del riesgo y escalando lo que funciona) lograrán una ventaja competitiva sostenible mediante una utilización óptima de los recursos y una automatización centrada en los resultados que crece con el éxito empresarial.

El futuro pertenece a las organizaciones que puedan combinar de forma inteligente la experiencia humana con las capacidades de la IA de los agentes. Esto crea modelos híbridos que ofrecen resultados superiores y, al mismo tiempo, mantienen la flexibilidad, la capacidad de aprendizaje y los beneficios de colaboración necesarios para las condiciones dinámicas del mercado.

Recursos

Los siguientes recursos pueden ayudarle a planificar, diseñar e implementar sistemas de inteligencia artificial para agencias en: AWS

- [Creación de arquitecturas sin servidor para la IA de los agentes \(orientación prescriptiva\)](#)
AWS
- [Cómo poner en práctica la IA de los agentes](#) (guía prescriptiva) AWS
- Los [patrones y flujos de trabajo de la IA de las agencias están en marcha \(orientación prescriptiva\)](#)
AWS
- [Inteligencia artificial \(orientación prescriptiva\)](#) AWS

- [Centro de optimización de costes de AWS](#) (Servicio de AWS)
- [Documentación de Amazon Bedrock](#) (Servicio de AWS)
- [Pilar de optimización de costes](#) (AWS Well-Architected Framework)
- [Agentes y soluciones de IA](#) (AWS Marketplace)

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	—	28 de enero de 2026

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:

- **Refactorizar/rediseñar:** traslade una aplicación y modifique su arquitectura mediante el máximo aprovechamiento de las características nativas en 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: Migrar la base de datos de Oracle en las instalaciones a Amazon Aurora PostgreSQL-Compatible Edition.
- **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: Migrar el sistema de administración de las relaciones con los clientes (CRM) a Salesforce.com.
- **Volver a alojar (migrar mediante lift-and-shift):** traslade una aplicación a la nube sin realizar cambios para aprovechar las capacidades 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

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.

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.

antipatrones

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 AIOps se utiliza en la estrategia de AWS migración, 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.

implementación azul/verde

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 más información, consulte el indicador [Implement break-glass procedures](#) en la guía de AWS Well-Architected.

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.

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: realizar inversiones fundamentales para escalar su adopción de la nube (p. ej., crear una landing zone, definir una CCoE, establecer un modelo de operaciones)

- Migración: migración de aplicaciones individuales
- Reinención: optimización de productos y servicios e innovación en la nube

Stephen Orban definió estas etapas en la entrada del blog [The Journey Toward Cloud-First & the Stages of Adoption en el 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 canalización de CI/CD 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 (CI/CD)

El proceso de automatización de las etapas de origen, compilación, 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 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.

defense-in-depth

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 defense-in-depth enfoque podría combinar la autenticación multifactorial, 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 ante desastres de cargas de trabajo en AWS: Recovery in the Cloud in the AWS Well-Architected Framework](#).

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. Este concepto lo introdujo Eric Evans en su libro, *Diseño impulsado por el dominio: abordando la complejidad en el corazón del 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 Microsoft ASP.NET \(ASMX\) 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. Los sistemas big-endianos almacenan primero el byte más significativo. Los sistemas Little-Endian 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 AWS PrivateLink y conceder permisos a otras Cuentas de AWS o a responsables AWS Identity and Access Management (de IAM). 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, mediante el que los modelos aprenden a partir de ejemplos (pasos) incrustados en las peticiones. La técnica de peticiones con pocos pasos puede ser eficaz para las 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)

Una gran red neuronal de aprendizaje profundo que se ha estado entrenando con conjuntos de datos masivos de datos generalizados y sin etiquetar. FMs son capaces de realizar 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?](#)

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 [la sección 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 el cumplimiento en todas las unidades organizativas (OUs). 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.

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.

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 de DevOps publicación típico.

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

IaC

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.

IloT

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 más información, consulte la práctica recomendada [Implementación mediante una infraestructura inmutable](#) en el Marco de 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 AWS de referencia de seguridad](#) recomienda configurar la cuenta de red con entradas, salidas e inspección VPCs para proteger la interfaz bidireccional entre la aplicación y el resto de Internet.

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

Término que introdujo [Klaus Schwab](#) en 2016 para referirse a la modernización de los procesos de fabricación mediante los avances en la conectividad, los datos en tiempo real, la automatización, el análisis, la IA y el 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 (T) Ilo

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 de la Internet de las cosas \(IIoT\) industrial](#).

VPC de inspección

En una arquitectura de AWS cuentas múltiples, una VPC centralizada que gestiona las inspecciones del tráfico de red VPCs entre Internet y las redes locales (en una misma o Regiones de AWS diferente). La [arquitectura AWS de referencia de seguridad](#) recomienda configurar su cuenta de red con entrada, salida e inspección VPCs para proteger la interfaz bidireccional entre la 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 obtener más información, consulte Qué son. LLMs](#)

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 para lo cual 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](#).

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 Framework.

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 machine-to-machine \(M2M\), basado en el patrón de publicación/suscripción, para dispositivos de IoT con recursos limitados.](#)

microservicio

Un servicio pequeño e independiente que se comunica a través de una red bien definida APIs 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 AWS servicios sin servidor](#).

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 un uso ligero. APIs 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

Equipos multifuncionales que agilizan la migración de las cargas de trabajo mediante enfoques automatizados y ágiles. Los equipos de las fábricas de migración suelen incluir a analistas y propietarios de operaciones, empresas, 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 fiabilidad y la previsibilidad, el AWS Well-Architected Framework 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](#).

Open Process Communications: arquitectura unificada (OPC-UA)

Un protocolo de machine-to-machine comunicación (M2M) para la automatización industrial. OPC-UA establece un estándar de interoperabilidad con esquemas de autenticación, autorización y cifrado 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 más información, consulte [Operational Readiness Reviews \(ORR\)](#) en el Marco de AWS Well-Architected.

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 para todos los miembros 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 AWS KMS (SSE-KMS) y las solicitudes dinámicas PUT y DELETE dirigidas 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 AWS de referencia de seguridad](#) recomienda configurar la cuenta de red con entradas, salidas e inspección VPCs para proteger la interfaz bidireccional entre la 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

Un contenedor que contiene información sobre cómo desea que Amazon Route 53 responda a las consultas de DNS de un dominio y sus subdominios dentro de uno o más VPCs. 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 la sección 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 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)

Política que proporciona un control centralizado de los permisos de todas las cuentas de una organización en AWS Organizations. SCPs defina barreras o establezca límites a las acciones que un administrador puede delegar en usuarios o roles. Puede utilizarlas SCPs como listas de permitidos o rechazados para especificar qué servicios o acciones están permitidos o prohibidos. Para obtener más información, consulte [las políticas de control de servicios](#) 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](#).

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](#).

split-and-seed modelo

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 con la aplicación de 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

Pares clave-valor que actúan como metadatos para organizar los recursos. AWS 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.

puerta de enlace de tránsito

Un centro de tránsito de red que puede usar para interconectar sus redes con VPCs las locales. 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

Una conexión entre dos VPCs que le 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.