



Economia per l'intelligenza artificiale agentica su AWS

AWS Guida prescrittiva



AWS Guida prescrittiva: Economia per l'intelligenza artificiale agentica su AWS

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

I marchi e l'immagine commerciale di Amazon non possono essere utilizzati in relazione a prodotti o servizi che non siano di Amazon, in una qualsiasi modalità che possa causare confusione tra i clienti o in una qualsiasi modalità che denigri o discrediti Amazon. Tutti gli altri marchi non di proprietà di Amazon sono di proprietà dei rispettivi proprietari, che possono o meno essere affiliati, collegati o sponsorizzati da Amazon.

Table of Contents

Introduzione	1
Destinatari principali	1
Obiettivi	2
Informazioni su questa serie di contenuti	2
Comprendere l'economia dell'IA agentic	3
Valutazione delle attività	3
Valutazione dell'impatto del rischio	4
Ritorno sull'investimento	6
Misurazione del successo e del ROI	7
Usa le tue fondamenta	7
Stabilisci obiettivi	7
Tieni traccia delle metriche	7
Usa AgentOps	8
Valutazione dei costi dei processi umani	8
Costi del lavoro	9
Costi prestazionali	9
Costi tecnologici	10
Costi legati alle opportunità	11
Costi relativi a rischi e difetti	11
Implementazione di sistemi di intelligenza artificiale agentic	12
Incorporare il feedback umano	13
Apprendimento comportamentale	13
Apprendimento continuo	14
Collaborazione uomo-IA	14
Prezzi basati sui risultati	15
Modello tradizionale e iniziale	15
Modello basato sui risultati	16
Usando Marketplace AWS	17
Caso di studio: operazioni di reclutamento	18
Scenario A	18
Struttura dei costi di base	19
Parametri operativi	20
Analisi dei costi basata sul volume	21
Analisi del ROI	21

Confronto dei costi cumulativi	22
Altri vantaggi	22
Scenario B	22
Struttura dei costi di base	23
Parametri operativi	24
Analisi dei costi basata sul volume	24
Analisi del ROI	25
Confronto dei costi cumulativi	25
Scenari a confronto	26
Conclusioni e risorse	28
Resources	28
Cronologia dei documenti	30
Glossario	31
#	31
A	32
B	35
C	37
D	40
E	44
F	46
G	48
H	49
I	50
L	53
M	54
O	58
P	61
Q	63
R	64
S	67
T	71
U	72
V	73
W	73
Z	74
.....	lxxvi

Economia per l'intelligenza artificiale agentica su AWS

Hans Schabert e Prasanta Roy, Amazon Web Services

Gennaio 2026 ([storia del documento](#))

Le organizzazioni che adottano sistemi di automazione basati sull'intelligenza artificiale e di intelligenza artificiale agentica devono prendere decisioni economiche informate tra lavoro umano e agenti intelligenti. Questo diventa fondamentale per operazioni cloud sostenibili. Questa guida ti aiuta a valutare, implementare e ottimizzare i compromessi economici tra la forza lavoro umana e i sistemi di intelligenza artificiale agentica su AWS. È possibile massimizzare il ritorno sull'investimento (ROI) mantenendo l'eccellenza operativa.

Nessun sistema è giusto al 100%. Questo principio fondamentale guida l'analisi economica dei sistemi di intelligenza artificiale umana e agentica. Le organizzazioni devono andare oltre i semplici confronti dei costi per valutare l'impatto economico totale, i profili di rischio, i requisiti di qualità delle decisioni e la creazione di valore strategico a lungo termine.

Il comportamento dei clienti sta cambiando drasticamente dai tradizionali investimenti tecnologici iniziali a pay-per-outcome modelli che allineano i costi ai risultati aziendali. Questa trasformazione richiede nuovi approcci per la valutazione, l'implementazione e l'ottimizzazione della collaborazione uomo-agente.

Il percorso verso il successo segue uno schema chiaro: inizia con lavori appropriati, misura tutto e amplia ciò che funziona. Organizzazioni che adottano questo approccio ottengono un vantaggio competitivo sostenibile attraverso l'allocazione intelligente delle risorse e l'automazione incentrata sui risultati.

Destinatari principali

Questa guida è destinata a quanto segue:

- Dirigenti (CEOs, CTOs, CFOs) che prendono decisioni strategiche in materia di investimenti
- Architetti aziendali che progettano strategie di automazione organizzativa
- Professionisti delle operazioni finanziarie che stanno ottimizzando la gestione finanziaria del cloud
- Leader tecnologici che stanno valutando gli approcci di implementazione dell'IA
- Responsabili delle unità aziendali che desiderano comprendere il ROI dell'automazione

- Professionisti del procurement che stanno esplorando nuovi modelli di prezzo basati sull'intelligenza artificiale

Per comprendere i concetti di questa guida, ti consigliamo di consultare [Foundations of agentic AI](#) on AWS

Obiettivi

Questa guida ti aiuta a comprendere quanto segue:

- Come valutare il potenziale dell'automazione degli agenti
- Modelli economici per confrontare il costo del lavoro umano con gli investimenti nei sistemi di intelligenza artificiale agentica
- Pay-per-outcome modelli di prezzo e loro impatto sull'economia dei progetti di intelligenza artificiale
- Tecniche di misurazione per dimostrare il ROI e gestire il rischio
- Strategie di scalabilità che trasformano i costi fissi in risultati variabili

Informazioni su questa serie di contenuti

Questa guida fa parte di una serie sull'intelligenza artificiale agentica su AWS. Per ulteriori informazioni e per visualizzare le altre guide di questa serie, consulta [Agentic AI](#) sul sito Web Prescriptive Guidance. AWS

Comprendere l'economia dell'IA agentica su AWS

Uno dei principi chiave è determinare quando utilizzare gli agenti di intelligenza artificiale e quando utilizzare i metodi deterministici tradizionali. Le organizzazioni devono valutare sistematicamente quali lavori giustificano l'automazione agentica e quali dovrebbero utilizzare l'automazione tradizionale o il funzionamento umano continuo. Questa decisione richiede la comprensione della relazione tra le caratteristiche dell'attività, la tolleranza al rischio e l'approccio operativo.

Prima di decidere di implementare l'intelligenza artificiale agentica, è necessario utilizzare il quadro decisionale per comprendere l'impatto economico. Il quadro decisionale include le seguenti tre domande chiave:

1. [Valutazione delle attività](#): questo compito è adatto a un agente di intelligenza artificiale?
2. [Valutazione dell'impatto del rischio](#): quali sono i rischi connessi?
3. [Ritorno sull'investimento](#): sarà conveniente?

Valutazione delle attività

Le attività con regole decisionali standardizzate e ad alta complessità possono trarre vantaggio dagli approcci di intelligenza artificiale agentica. Le attività semplici e altamente standardizzate sono meglio soddisfatte dall'automazione tradizionale o dall'automazione robotica dei processi. I sistemi di intelligenza artificiale agentica eccellono nel ragionamento, nella comprensione del contesto o nel prendere decisioni in modo adattivo. Agentica aggiungono valore oltre l'elaborazione basata su regole. Le implementazioni di intelligenza artificiale agentica di successo richiedono sistemi in grado di apprendere e adattarsi.

Quando valuti un'attività, considera i seguenti fattori:

- **Complessità**: è richiesto il grado di ragionamento e la comprensione del contesto. Le attività che richiedono una comprensione contestuale, un'interpretazione dettagliata o risposte adattive alle mutevoli condizioni favoriscono gli approcci agentici rispetto all'automazione tradizionale, mentre le attività puramente meccaniche o di calcolo potrebbero non richiedere l'intelligenza agentica.
- **Standardizzazione**: presenza di schemi e regole chiari. L'intelligenza artificiale agentica è consigliata se l'attività richiede una comprensione contestuale. Se non è necessario alcun adattamento o apprendimento, prendi in considerazione l'automazione tradizionale.

- **Volume:** frequenza delle prestazioni delle attività. L'intelligenza artificiale agentica è consigliata per attività autonome. L'automazione tradizionale è consigliata per attività consistenti e ad alto volume. Tuttavia, il volume da solo non determina l'approccio. Decisioni di basso volume e di alto valore potrebbero giustificare l'assistenza degli agenti per una migliore qualità delle decisioni piuttosto che per la riduzione dei costi.
- **Valore:** impatto aziendale per completamento delle attività. Prendi in considerazione l'intelligenza artificiale agentica per risultati di alto valore che richiedono capacità autonome simili a quelle umane. Prendi in considerazione l'automazione tradizionale per attività ripetute e coerenti, che possono essere eseguite in modo deterministico.

Valutazione dell'impatto del rischio

Attualmente esistono quattro approcci di implementazione dell'intelligenza artificiale agentica: completamente autonomo, «human in the loop», copilota o guidato dall'uomo con supporto agli agenti. Ciascuno ha il proprio profilo di rischio e la propria tolleranza agli errori, e tutti coinvolgono in qualche modo gli esseri umani. La tabella seguente descrive i dettagli del rischio di questi approcci.

Livello di autonomia	Profilo di rischio	Tolleranza agli errori	Esempi di casi d'uso	Coinvolgimento umano
Completamente autonomo	Basso rischio	1-2% accettabile	<ul style="list-style-type: none"> • Categorizzazione dei dati di base • Instradamento dei documenti • Generazione di report standard 	<ul style="list-style-type: none"> • Supervisione minima • Audit periodici
Umano al centro	Rischio medio	Inferiore allo 0,5%	<ul style="list-style-type: none"> • Bozze di risposte 	<ul style="list-style-type: none"> • Revisione regolare

			<ul style="list-style-type: none"> • Moderazione dei contenuti • Elaborazione iniziale dei reclami 	<ul style="list-style-type: none"> • Gestione dell'eccezione • Garanzia di qualità
Copilota	Rischio elevato	Quasi zero	<ul style="list-style-type: none"> • Input di pianificazione strategica • Valutazioni del rischio • Decisioni di investimento 	<ul style="list-style-type: none"> • L'essere umano prende le decisioni finali • L'agente fornisce consigli
Guidato da un operatore umano con supporto da parte degli agenti	Rischio critico	Tolleranza zero	<ul style="list-style-type: none"> • Decisioni legali • Diagnosi medica • Conformità alle normative 	<ul style="list-style-type: none"> • Il processo Human Drive • L'agente fornisce solo informazioni di ricerca o analisi e di supporto

La tabella seguente descrive le considerazioni chiave nella scelta tra questi approcci.

Considerazione	Completamente autonomo	Un essere umano al centro	Copilota	Guidato dall'uomo
Efficienza dei costi	Massimo	Elevata	Media	Bassa

Scalabilità	Illimitato	Elevata	Media	Limitato
Velocità di elaborazione	La più veloce	Veloce	Media	Lento
Gestione del rischio	Di base	Migliorato	Forte	Il più forte
Gestione della complessità	Compiti semplici	Compiti moderatamente complessi	Compiti complessi	Attività critiche

Questo framework di considerazione aiuta le organizzazioni ad abbinare i livelli di autonomia ai profili di rischio, scalare le operazioni in modo appropriato, bilanciare efficienza e controllo, implementare una governance adeguata e ottimizzare l'allocazione delle risorse.

Ritorno sull'investimento

Il calcolo del ritorno sull'investimento per i sistemi di intelligenza artificiale agentica inizia con un'analisi completa dei costi. Le organizzazioni devono innanzitutto calcolare i costi umani correnti, inclusi salari, benefit e spese relative all'area di lavoro, oltre alle spese specifiche dei processi e ai costi nascosti come formazione, copertura e tempi di inattività.

Per un'analisi del pareggio, le organizzazioni devono considerare i costi di implementazione, le spese operative correnti e il volume necessario per giustificare l'investimento. È inoltre importante tenere conto delle variazioni stagionali e dei vantaggi della curva di apprendimento che emergono man mano che i sistemi maturano e migliorano nel tempo.

Nel valutare gli agenti di intelligenza artificiale, le organizzazioni devono ricordare che questi sistemi in genere hanno costi iniziali più elevati ma costi per transazione inferiori rispetto alle operazioni umane. Inoltre, gli agenti di intelligenza artificiale dimostrano di migliorare le prestazioni nel tempo e offrono una migliore scalabilità rispetto ai team umani. Ciò li rende sempre più convenienti man mano che l'implementazione è scalabile e l'esperienza operativa accumulata.

Misurazione del successo e del ROI dei sistemi di intelligenza artificiale agentica

Misurare il successo nell'implementazione di sistemi di intelligenza artificiale agentici richiede un approccio sistematico. Questa sezione fornisce una metodologia chiara per la valutazione e l'ottimizzazione continua che utilizza l'analisi esistente anziché partire da zero.

Fase 1: Usa le fondamenta esistenti

Iniziate con una valutazione completa dei costi in base ai consigli contenuti nella sezione [Valutazione dei costi di processo correnti](#). Ciò fornisce una base operativa per i calcoli del ROI. Come descritto nella sezione [Valutazione dell'impatto del rischio](#), scegliete tra i quattro livelli di autonomia (completamente autonomo, human in loop, approccio copilota, guida umana con supporto degli agenti) per determinare i criteri di misurazione e le soglie di tolleranza agli errori appropriati per ciascun processo.

Fase 2: Stabilisci obiettivi di successo chiari

Stabilisci l'architettura e gli obiettivi di successo che enfatizzino i sistemi in grado di apprendere, come descritto nella sezione [Modelli di successo per l'implementazione di sistemi di intelligenza artificiale agentica](#). Concentrati sul miglioramento continuo piuttosto che sulle prestazioni statiche. Stabilisci le tempistiche del ROI utilizzando la metodologia di analisi del pareggio illustrata nel [Case study: Comparazione dei costi dell'IA umana e degli agenti per le operazioni](#) di reclutamento. Includi punti decisionali chiari per licenziare gli agenti non performanti.

Fase 3: Tieni traccia delle metriche chiave

Monitora le prestazioni finanziarie rispetto alla linea di base stabilita e monitora i risparmi sui costi e i miglioramenti strategici del valore. Misura le metriche operative, compresi i tassi di errore entro soglie accettabili per il livello di autonomia prescelto, i miglioramenti della velocità di elaborazione e gli incrementi di coerenza. Concentrati su indicatori strategici che dimostrino la capacità di apprendimento e l'adattamento nel tempo.

Fase 4: Utilizzo AgentOps

Applica il framework di apprendimento continuo della sezione [Incorporazione del feedback umano nei sistemi di intelligenza artificiale agentici](#) per ottimizzare il processo decisionale attraverso l'integrazione sistematica del feedback umano. Crea sistemi di apprendimento in tempo reale che incorporino le conoscenze umane per il miglioramento delle prestazioni. Monitora la trasformazione verso modelli di business basati sui risultati come descritto in [Trasformazione economica verso prezzi basati sui risultati per i sistemi di intelligenza artificiale agentica](#) su AWS.

Valutazione dei costi attuali dei processi umani

Comprendere i costi reali dei processi è fondamentale per prendere decisioni informate sugli investimenti nei sistemi di intelligenza artificiale agentica. Innanzitutto, è necessario stabilire una base precisa del costo attuale dei processi, comprese tutte le spese nascoste, i tassi di fallimento e i costi legati alle opportunità. Ciò consente di sviluppare calcoli precisi del ROI e prendere decisioni strategiche. Questa valutazione completa dei costi funge da base fondamentale per valutare se i sistemi di intelligenza artificiale agentica possono offrire un valore reale come compagni produttivi.

La valutazione dei costi di base è essenziale per i seguenti motivi principali:

- **Precisione del ROI:** baseline di costo accurate supportano proiezioni realistiche del ROI che tengono conto dell'intero spettro delle spese operative correnti.
- **Strategia di implementazione degli agenti:** una comprensione completa dei costi aiuta le organizzazioni a identificare i processi più promettenti per l'implementazione iniziale dei sistemi di intelligenza artificiale agentica.
- **Misurazione delle prestazioni:** le linee di base consolidate forniscono il quadro di misurazione per tracciare i vantaggi effettivi e previsti delle implementazioni di intelligenza artificiale agentica.

Le organizzazioni devono identificare e valutare sistematicamente tutti i fattori di costo che influenzano l'economia dei processi prima di confrontare alternative umane e agenti. Questa valutazione garantisce calcoli di base accurati tenendo conto dei fattori di costo evidenti e nascosti. Pone particolare enfasi sui costi di fallimento, sui tassi storici di fallimento e sulle opportunità commerciali mancate, che rappresentano il vero costo totale dei processi correnti.

Questa sezione descrive come raccogliere dati in ciascuna categoria di costo per stabilire misurazioni di base accurate per i processi correnti. Descrive le fonti di informazione e fornisce esempi per le seguenti categorie di costi:

- [Costi del lavoro](#)
- [Costi di coerenza e prestazioni umane](#)
- [Costi tecnologici e infrastrutturali](#)
- [Costi legati alle opportunità commerciali perse](#)
- [Costi relativi a rischi e difetti](#)

Costi del lavoro

Estrai 24 mesi di dati sulle retribuzioni che includono lo stipendio base, gli straordinari, i benefit e i costi di formazione. Utilizza il tuo sistema informativo sulle risorse umane (HRIS) per tenere traccia delle spese di assunzione e dei tassi di turnover. I sistemi di monitoraggio del tempo rivelano la produttività effettiva rispetto alle ore programmate. Le piattaforme di gestione delle prestazioni mostrano la correlazione tra i livelli di competenza e i costi di compensazione. Calcola le tariffe orarie a pieno carico da destinare ai costi generali di gestione.

Di seguito è riportato un elenco di esempio dei fattori di costo della manodopera.

Fattore di costo	Impatto aziendale
Compensazione di base	25-150 USD all'ora a pieno carico
Vantaggi e imposte sui salari	25-40% dello stipendio base
Formazione e sviluppo	5-15% del costo annuo della manodopera
Sovraccarico di gestione	15-25% del costo diretto della manodopera

Costi di costanza e prestazioni umane

Combina i dati dei sistemi di gestione dei progetti che mostrano le variazioni nel completamento delle attività con i sistemi di rilevazione delle presenze. Ciò può rivelare modelli di assenteismo e cambiamenti stagionali. Le piattaforme di assistenza clienti dimostrano gli intervalli di prestazioni individuali attraverso metriche di risoluzione, mentre i dati relativi alle vendite e alla gestione delle relazioni con i clienti (CRM) possono mostrare variazioni di efficienza nella chiusura delle trattative. I sistemi di gestione della qualità forniscono tassi di difetto e dati sulla conformità dei processi tra team e sedi. I sistemi di workflow registrano i tempi di completamento, i ritardi di approvazione

e la frequenza di gestione delle eccezioni. L'analisi della comunicazione rivela il sovraccarico di coordinamento dovuto alla frequenza delle riunioni e ai modelli di collaborazione.

Di seguito è riportato un elenco di esempi di fattori di costo per le prestazioni e la coerenza umane.

Fattore di costo	Impatto aziendale
Fluttuazioni della produttività	Intervallo di prestazioni del 20-50%
Assenteismo e copertura	È necessaria una capacità aggiuntiva del 15-25%
Cicli di affaticamento e motivazione	Variazione della produttività del 10-30%
Incoerenze nella procedura	Perdita di efficienza del 10-40%
Variazioni del controllo di qualità	10-30% del costo totale
Coordinamento generale	15-25% dei costi operativi

Costi tecnologici e infrastrutturali

Le piattaforme di gestione delle licenze mostrano i costi del software e i tassi di utilizzo. Il monitoraggio dell'infrastruttura fornisce dati sull'operatività, metriche delle prestazioni e costi di manutenzione. I sistemi di help desk tengono traccia dei costi generali del supporto e dei problemi tecnici ricorrenti. I sistemi di gestione dei fornitori tengono conto dei costi totali delle relazioni tecnologiche, comprese le spese di integrazione e le prestazioni a livello di servizio.

Di seguito è riportato un elenco di esempi di fattori di costo per la tecnologia e l'infrastruttura.

Fattore di costo	Impatto aziendale
Sistemi tecnologici	\$50-500 per utente al mese
Spazio di lavoro e attrezzature	200-1000 dollari per dipendente al mese

Costi legati alle opportunità commerciali perse

Le piattaforme CRM contengono i tempi di risposta dei lead, i tassi di conversione e la documentazione relativa alle opportunità perse. L'automazione del marketing mostra l'impatto dei ritardi di follow-up sulle conversioni di lead. I sistemi di assistenza clienti rivelano in che modo i problemi operativi influiscono sulla soddisfazione e sulla fidelizzazione. L'analisi della concorrenza fornisce i requisiti di risposta del mercato e dati relativi alle vittorie o alle perdite che collegano le prestazioni operative ai risultati di fatturato.

Di seguito è riportato un elenco di esempi di fattori di costo che determinano la perdita di opportunità commerciali.

Fattore di costo	Impatto aziendale
Ritardi nella risposta del mercato	Entrate per giorno di ritardo
Limiti di capacità	Opportunità commerciali perse
Allocazione delle risorse per l'innovazione	Costo di opportunità del lavoro di routine
Ritardi nell'acquisizione dei clienti	Perdita di piombo del 50-90% a causa della risposta lenta

Costi relativi a rischi e difetti

La documentazione della polizza assicurativa mostra i costi per la responsabilità generale, la responsabilità professionale, la compensazione dei lavoratori e la copertura della responsabilità informatica. I rapporti interni di valutazione del rischio identificano le vulnerabilità operative e i costi di mitigazione associati. I sistemi di tracciamento dei difetti documentano i guasti dei prodotti o dei servizi, inclusi i costi di rilevamento, le spese di sostituzione e le richieste di garanzia. I programmi di sostituzione degli asset mostrano i tassi di guasto delle apparecchiature e i costi di sostituzione. I report sugli incidenti di sicurezza tengono traccia degli infortuni sul lavoro e delle relative richieste di risarcimento dei lavoratori. I piani di continuità operativa descrivono in dettaglio i costi dei sistemi di backup e gli investimenti per il disaster recovery.

Di seguito è riportato un elenco di esempi di fattori di costo relativi a rischi e difetti.

Fattore di costo	Impatto aziendale
------------------	-------------------

Costi assicurativi	1-5% del budget operativo
Costo degli errori	50-5.000 USD per incidente di errore
Impatto dell'errore umano	2-15% del costo operativo totale
Tassi di errore e rilavorazioni	1,5-4 volte il costo originario delle correzioni

Modelli di successo per l'implementazione di sistemi di intelligenza artificiale agentica su AWS

[State of Enterprise AI Adoption](#) (rapporto ISG 2025) rivela che il principale ostacolo all'implementazione efficace dell'IA non è la capacità tecnica ma il divario di apprendimento. Questo termine si riferisce a sistemi che non possono adattarsi, ricordare il contesto o migliorare nel tempo. Organizations che implementano strumenti di intelligenza artificiale statici registrano tassi di fallimento elevati. Di seguito sono riportate le caratteristiche comuni dei sistemi di intelligenza artificiale agentici che raggiungono il successo:

- Memoria contestuale: sistemi che conservano la cronologia delle conversazioni e le preferenze dell'utente
- Integrazione del feedback: capacità di imparare dalle correzioni e migliorare le prestazioni
- Adattamento del flusso di lavoro: adattamento automatico alle mutevoli esigenze aziendali
- Miglioramento continuo: miglioramento misurabile attraverso l'esperienza operativa

Organizations che ottengono implementazioni AI di successo spesso danno la priorità a quanto segue:

- Utilizzare ecosistemi di partner completi anziché creare ed esplorare in modo indipendente le funzionalità di intelligenza artificiale
- Sistemi in grado di apprendere rispetto a strumenti statici
- Concentrazione dei risultati aziendali rispetto al confronto delle caratteristiche tecniche
- Integrazione del flusso di lavoro anziché strumenti autonomi
- Adattamento continuo anziché implementazione una tantum

[Questi modelli sono in linea con molte Servizio AWS funzionalità, in particolare l'accesso al modello di base in Amazon Bedrock, l'architettura basata sugli eventi e il monitoraggio AWS Lambda completo offerto tramite Amazon CloudWatch](#) Per ulteriori informazioni sull'integrazione del feedback umano e dei sistemi in grado di apprendere, consulta la sezione [Incorporazione del feedback umano](#) nei sistemi di intelligenza artificiale agentici di questa guida.

Incorporazione del feedback umano nei sistemi di intelligenza artificiale agentica

Nessun sistema ha successo al 100% e il fallimento è inevitabile. A ogni fallimento corrisponde un costo di cambiamento. Human in the loop è un approccio di intelligenza artificiale in cui l'IA svolge un'attività, ma è necessario l'intervento o l'approvazione umana. Questo approccio deve essere utilizzato quando il costo del fallimento è superiore al costo di una human-in-the-loop soluzione.

Il successo dei sistemi di intelligenza artificiale agentica dipende fundamentalmente dalla capacità dell'operatore di apprendere e migliorare attraverso il feedback umano. Il costo dello sforzo umano deve essere preso in considerazione, a seconda del livello di impegno richiesto. A differenza degli strumenti di automazione statici che eseguono regole predeterminate, le human-in-the-loop soluzioni dispongono di sistemi agentici in grado di apprendere che creano una partnership dinamica tra agenti autonomi e umani. L'esperienza umana migliora continuamente le prestazioni dell'agente, mentre gli agenti gestiscono l'elaborazione di routine su larga scala. Questo approccio collaborativo trasforma l'implementazione dell'IA da un'implementazione una tantum a un processo di ottimizzazione continuo. Il sistema si adatta ai modelli organizzativi, internalizza gli standard di qualità e affina le sue capacità decisionali sulla base dell'esperienza operativa del mondo reale. Acquisendo sistematicamente le correzioni, le approvazioni e le informazioni umane, le organizzazioni possono creare agenti di intelligenza artificiale in grado di comprendere il contesto, riconoscere i modelli e allinearsi sempre più agli obiettivi aziendali nel tempo.

Per le soluzioni che non richiedono l'intervento o il supporto umano, non è necessario tenere conto dei costi specifici dell'uomo nell'economia degli agenti.

Apprendimento comportamentale da operatori umani

Gli operatori umani forniscono un feedback fondamentale che i sistemi di intelligenza artificiale agentica possono utilizzare per apprendere, adattare e migliorare le loro risposte nel tempo. Questo ciclo di feedback crea un ambiente collaborativo in cui l'esperienza umana migliora le capacità degli agenti mentre gli agenti gestiscono l'elaborazione di routine.

Attraverso il riconoscimento dei modelli di comportamento umano, gli agenti imparano dai modelli di interazione umana per rispecchiare approcci di comunicazione efficaci. Questo li aiuta ad adattarsi ai modelli decisionali organizzativi e ai livelli di tolleranza al rischio. I sistemi interiorizzano le aspettative di qualità attraverso correzioni e approvazioni umane. Possono anche apprendere le risposte appropriate per diversi segmenti di clienti e contesti aziendali.

Meccanismi efficaci di raccolta del feedback acquisiscono sistematicamente le modifiche e le modifiche apportate dall'operatore alle risposte degli agenti. Analizzano ciò che i revisori umani approvano, rifiutano o modificano nelle raccomandazioni degli agenti. Comprendendo perché determinati casi richiedono l'intervento umano e incorporando la valutazione umana delle prestazioni degli agenti in diversi scenari e livelli di complessità, questi sistemi perfezionano continuamente le proprie capacità per allinearsi maggiormente agli standard e alle aspettative organizzative.

Operazioni di apprendimento continuo

L'integrazione dell'apprendimento in tempo reale consente ai sistemi di intelligenza artificiale agentici di incorporare il feedback umano e migliorare immediatamente le risposte degli agenti attraverso l'aggiornamento dinamico del modello. Questi sistemi utilizzano le conoscenze umane per identificare nuovi modelli e casi limite. Ciò migliora le loro capacità di riconoscimento dei modelli mentre costruiscono la memoria organizzativa attraverso esperienze di apprendimento guidate dall'uomo. Il perfezionamento continuo basato sul feedback degli operatori umani e sui risultati aziendali favorisce l'ottimizzazione continua delle prestazioni.

La formazione guidata da persone acquisisce le conoscenze degli esperti per migliorare le capacità decisionali degli agenti. Trasferisce le competenze fondamentali degli operatori esperti al sistema di intelligenza artificiale. Attraverso l'apprendimento basato su scenari, i sistemi utilizzano esempi creati dall'uomo per migliorare la gestione di situazioni complesse. Inoltre, allineano gli standard prestazionali degli agenti alle aspettative di qualità umana attraverso la calibrazione della qualità. Questo approccio incorpora le conoscenze umane sulla cultura organizzativa e sulle aspettative dei clienti. Questo adattamento culturale aiuta gli agenti a rispondere in modo appropriato in diversi contesti.

Eccellenza operativa con la collaborazione uomo-intelligenza artificiale

L'ottimizzazione automatizzata basata sulla consapevolezza del rischio consente la valutazione continua delle condizioni operative e della probabilità di errore con la supervisione umana per scenari ad alto rischio. Questo aiuta i sistemi a imparare dalle valutazioni del rischio umano e a migliorare i processi decisionali futuri. [Amazon Bedrock](#) fornisce l'accesso a più modelli di base con funzionalità e profili di costo diversi. Ciò consente un routing intelligente che considera sia i profili di costo che

di rischio, incorporando al contempo il feedback umano per ottimizzare la selezione dei modelli. L'ottimizzazione delle prestazioni bilancia l'efficienza con la riduzione al minimo del tasso di errore incorporando il feedback umano sugli standard di qualità e sui compromessi accettabili in termini di prestazioni. Le decisioni automatizzate prendono in considerazione il costo totale di proprietà adeguato al rischio. Gli operatori forniscono indicazioni sulla tolleranza al rischio organizzativo e sulla ponderazione delle priorità aziendali. Ciò consente di ottimizzare i costi allineandosi allo stesso tempo agli obiettivi organizzativi.

I sistemi di apprendimento potenziati dall'uomo danno priorità all'input umano in base all'impatto degli errori e alle conseguenze aziendali. In questo modo vengono creati sistemi di apprendimento in grado di comprendere sia l'accuratezza tecnica che il contesto aziendale attraverso feedback ponderati in base al rischio. L'analisi periodica delle prestazioni incorpora metriche di rischio e analisi dei costi degli errori, con informazioni umane che forniscono un contesto che i sistemi automatizzati non sono in grado di acquisire. Lo sviluppo delle migliori pratiche enfatizza la gestione del rischio e la prevenzione degli errori combinando il riconoscimento automatico dei pattern con l'esperienza e il giudizio umani. Lo sviluppo delle capacità organizzative attraverso programmi di formazione sviluppa sia le competenze umane per la gestione dei sistemi di intelligenza artificiale agentica sia le capacità degli agenti per supportare il processo decisionale umano. Ciò garantisce un approccio globale alla collaborazione uomo-IA che rafforza entrambe le componenti della partnership.

Trasformazione economica verso una tariffazione basata sui risultati per i sistemi di intelligenza artificiale agentica su AWS

Il passaggio dai tradizionali modelli a costo fisso a quelli basati sui risultati rappresenta una trasformazione fondamentale nel modo in cui le organizzazioni strutturano le proprie operazioni economiche e gestiscono il rischio. Questa trasformazione funge da percorso per la modernizzazione costante dei processi esistenti, finanziando al contempo la trasformazione dell'intelligenza artificiale agentica. Consente alle organizzazioni di evolversi da operazioni statiche e ad alta intensità di risorse a modelli di business dinamici e orientati ai risultati.

Modello tradizionale e iniziale

I reparti spesso operano come centri di costo con costi di manodopera diretti finanziati dall'allocazione dei costi. Le organizzazioni in genere desiderano ridurre questa allocazione dei costi. Se il processo non viene modernizzato, il reparto deve fornire gli stessi risultati con una forza lavoro più piccola. Questo in genere riduce la qualità. I modelli di business tradizionali creano sfide significative, tra cui:

- Scalabilità lineare dei costi in base all'aumento dei volumi: ciò richiede alle organizzazioni di assumere personale aggiuntivo per gestire l'aumento dei volumi.
- Impegni a costi fissi: questi impegni persistono indipendentemente dalle prestazioni aziendali e dall'efficienza dei processi.
- Pianificazione avanzata: la flessibilità limitata durante le recessioni economiche e i vincoli di capacità richiede una pianificazione anticipata.
- Ciclo di degrado della qualità: la riduzione dei budget porta a una diminuzione della qualità del servizio quando i costi vengono ridotti senza miglioramenti dei processi.

Modello basato sui risultati

I modelli moderni basati sui risultati legano i pagamenti direttamente a risultati aziendali misurabili, come assunzioni completate con successo, metriche di qualità raggiunte, miglioramenti dell'efficienza dei processi o incrementi di produttività realizzati. Ciò sposta fundamentalmente il rischio finanziario dalle unità aziendali ai fornitori di servizi, creando al contempo un allineamento naturale degli incentivi. Di seguito sono riportati i principali vantaggi di un modello basato sui risultati:

- I costi dipendono direttamente dal valore aziendale generato
- Allineamento naturale tra spese operative e ricavi
- Flessibilità di adeguare la capacità in base alle condizioni di mercato
- Pay-per-success i modelli riducono il rischio finanziario spostando l'esposizione finanziaria dall'investimento iniziale alla performance operativa continua
- Concentrati su sistemi in grado di apprendere e che migliorano nel tempo, piuttosto che su alternative statiche

Questa trasformazione va ben oltre i centri di costo interni e ridefinisce radicalmente il modo in cui le organizzazioni interagiscono con partner e fornitori di servizi esterni. Applicando prezzi basati sui risultati alle collaborazioni con i partner, le organizzazioni possono promuovere miglioramenti della qualità a lungo termine e ridurre i costi, enfatizzando indirettamente la modernizzazione dell'intelligenza artificiale agentica.

Organizations può sperimentare rapidamente, misurare chiaramente le prestazioni e scalare in base al valore aziendale effettivo generato anziché ai tradizionali impegni di risorse fisse. Questo approccio consente quanto segue:

- Evoluzione delle relazioni con i fornitori: i partner investono nel successo dei clienti piuttosto che nella semplice fornitura di servizi.
- Metriche di risultato standardizzate: semplifica i processi di approvvigionamento tra più fornitori.
- Reattività al mercato: adattamento rapido alle mutevoli condizioni di mercato e alle esigenze dei clienti.
- Vantaggio competitivo: utilizzo superiore delle risorse e capacità operative avanzate.
- Partnership orientate alla qualità: la collaborazione a lungo termine si concentra sul miglioramento continuo e su risultati misurabili.

Marketplace AWS Utilizzo come pay-per-outcome fattore abilitante

Il fattore chiave di questa trasformazione è [Marketplace AWS](#) che funge da veicolo di transazione per il lavoro degli agenti e la determinazione dei prezzi basata sui risultati. Fornisce l'accesso a centinaia di agenti di intelligenza artificiale e soluzioni agentiche preconfigurati con modelli di prezzo trasparenti e basati sull'utilizzo. Può aiutare a eliminare i costi iniziali delle licenze, ridurre la complessità dell'implementazione e consentire alle organizzazioni di concentrarsi su sistemi in grado di apprendere che si adattano e migliorano nel tempo piuttosto che su alternative statiche

L'utilizzo Marketplace AWS può offrire i seguenti vantaggi:

- Sperimentazione rapida: testa più soluzioni senza investimenti di capitale significativi
- Prezzi trasparenti: costi basati sull'utilizzo con chiara attribuzione ai risultati aziendali
- Soluzioni comprovate: accesso a agenti collaudati e forniti da fornitori esperti
- Integrazione integrata: connettività senza interruzioni con i sistemi esistenti Servizi AWS
- Riduzione del rischio: possibilità di cambiare fornitore in base alle prestazioni
- Accesso alle capacità di apprendimento: disponibilità di sistemi adattivi senza costi di sviluppo interni

Questo approccio consente alle organizzazioni di confrontare diverse opzioni in base alla fornitura dei risultati e alle capacità di apprendimento anziché agli elenchi di funzionalità. Può anche aiutarvi a stabilire criteri di successo e metodologie di misurazione chiari e a negoziare prezzi basati sui risultati legati ai risultati aziendali e al miglioramento del sistema. Finanziando la trasformazione agentica dell'IA attraverso modelli basati sui risultati, le organizzazioni possono modernizzare continuamente i propri processi pagando solo per miglioramenti misurabili e risultati di successo.

Caso di studio: confronto dei costi dell'IA umana e agentic per le operazioni di reclutamento

Le operazioni di reclutamento forniscono un case study convincente per valutare i compromessi economici tra sistemi di intelligenza artificiale umana e agentic, ma il calcolo del ROI dipende in modo determinante dall'attuale base operativa. Le organizzazioni che valutano gli investimenti agentic AI spesso si pongono una domanda fondamentale: «E se invece ottimizzassimo semplicemente i nostri processi umani esistenti?» Per affrontare direttamente questo problema, questa analisi presenta due scenari distinti che racchiudono la gamma di efficienza operativa umana.

[Lo scenario A](#) modella il curriculum vitae (CV) di 45 minuti o riprende i tempi di screening. [Lo scenario B](#) dimostra l'ottimizzazione delle operazioni umane alla velocità di 15 minuti per applicazione, con un miglioramento dell'efficienza del 66%. Ad esempio, questo miglioramento potrebbe essere ottenuto attraverso processi semplificati, reclutatori esperti o strumenti specializzati.

Confrontando funzionalità identiche dei sistemi di agenti con queste diverse linee di base relative alle prestazioni umane, riveliamo come l'efficienza dei processi esistenti influenzi i calcoli del ROI, le tempistiche di pareggio e le decisioni strategiche di implementazione. Questo approccio a doppio scenario ha molteplici scopi. Impedisce alle organizzazioni di abbandonare l'intelligenza artificiale agentic supponendo che la sola ottimizzazione dei processi sia sufficiente. Inoltre, aiuta le organizzazioni con processi già efficienti a comprendere le proprie caratteristiche economiche specifiche. Inoltre, questi scenari evidenziano quando i vantaggi non finanziari, come la disponibilità e la scalabilità 24 ore su 24, 7 giorni su 7, diventano fattori decisionali primari. La comprensione di queste dinamiche economiche attraverso diverse linee di base di efficienza consente alle organizzazioni di prendere decisioni informate su dove e quando implementare sistemi di intelligenza artificiale agentic per il massimo impatto aziendale.

Scenario A: tempo di screening di 45 minuti

Lo scenario A rappresenta le operazioni di reclutamento in cui i reclutatori umani dedicano 45 minuti allo screening di ogni curriculum. Questo scenario modella un recruiter di medio livello con un costo annuo a pieno carico di 112.250 dollari. Questo addetto alla selezione del personale elabora le candidature durante le normali ore lavorative con caratteristiche tipiche delle prestazioni umane. Al contrario, il sistema di intelligenza artificiale agentic richiede un investimento iniziale di 23.000 dollari per lo sviluppo, la personalizzazione e l'integrazione ATS e ha un costo operativo mensile minimo di 500 dollari per l'infrastruttura cloud. L'agente elabora le applicazioni in soli 5 minuti con

disponibilità 24 ore su 24, 7 giorni su 7, raggiungendo un tasso di errore del 2% e una capacità mensile superiore a 8.600 applicazioni. Si tratta di un notevole divario di efficienza, in cui l'agente opera 9 volte più velocemente per applicazione e una capacità mensile 39 volte maggiore. Questa sezione esamina l'analisi della struttura dei costi, le metriche operative, i confronti basati sui volumi e i calcoli cumulativi del ROI nei primi sei mesi di attività.

Struttura dei costi di base

La tabella seguente mostra i costi di configurazione iniziali per lo scenario A.

Componente	Operazioni umane	Sistema di intelligenza artificiale agentica
Sviluppo e personalizzazione degli agenti	N/D	15.000\$
Integrazione del sistema di tracciamento dei richiedenti (ATS)	N/D	5.000\$
Formazione e ottimizzazione	N/D	3.000 dollari
Costo totale di configurazione iniziale	\$0	23.000\$

La tabella seguente mostra i costi fissi annuali per lo scenario A.

Componente	Operazioni umane	Sistema di intelligenza artificiale agentica
Salario base	65.000 dollari	N/D
Vantaggi (30%)	19.500\$	N/D
Spazio di lavoro e attrezzature	12.000\$	N/D
Supervisione della gestione (15%)	9.750\$	N/D

Formazione e sviluppo	6.000 dollari	N/D
Costo fisso annuo totale	112.250\$	N/D

La tabella seguente mostra i costi operativi mensili per lo scenario A.

Componente	Operazioni umane	Sistema di intelligenza artificiale agentic
Cloud computing	N/D	\$200
Archiviazione	N/D	\$100
operazioni database	N/D	\$100
Monitoraggio	N/D	\$100
Costo fisso mensile totale	9.354\$	\$500

Parametri operativi

La tabella seguente mostra le metriche operative per lo scenario A.

Parametro	Operazioni umane	Sistema di intelligenza artificiale agentic
Tempo di elaborazione per applicazione	45 minuti	5 minuti
Capacità oraria	1.3 Applicazioni	12 applicazioni
Capacità giornaliera (24 ore)	10-11 applicazioni	288 applicazioni
Capacità mensile	220 applicazioni	8.640 applicazioni
Costo per applicazione	\$45	\$2,50

Costo per assunzione andata a buon fine	2.200\$	\$125
Tasso di errore	5%	2%
Costo di correzione degli errori	90 USD per errore	45 dollari per escalation

Analisi dei costi basata sul volume

La tabella seguente mostra un'analisi dei costi basata sul volume per lo scenario A. In questo esempio, il costo del sistema Agentic AI include i costi fissi e i costi di configurazione ammortizzati di 1.917 dollari al mese per 12 mesi.

Volume mensile	Costo umano	Costo del sistema Agentic AI	Risparmi mensili
100 applicazioni	4.500\$	750 dollari	3.750 dollari
500 applicazioni	22.500\$	2.667\$	\$19.833
1.000 applicazioni	45.000\$	4.917\$	\$40.083

Analisi del ROI

La tabella seguente mostra un'analisi del ROI per lo scenario A basata sull'elaborazione di 500 applicazioni al mese.

Parametro	Valore
Costo umano mensile	22.500 dollari
Costo mensile dell'agente	2.667\$
Risparmi mensili	19.833\$
Risparmio annuo	\$237.996
Periodo di pareggio	1,16 mesi

Confronto dei costi cumulativi

La tabella seguente mostra un confronto dei costi cumulativi per lo scenario A per i primi sei mesi, ipotizzando 500 domande al mese.

Mese	Costo umano	Costo del sistema Agentic AI	Risparmi cumulativi
1	22.500\$	\$25,667	- 3.167\$
2	45.000\$	\$28.334	16.666\$
3	67.500\$	31.001\$	\$36.499
4	90.000\$	\$33,668	\$56,332
5	\$112.500	\$36.335	76,165\$
6	135.000\$	\$39.002	\$95,998

Ulteriori vantaggi del sistema di intelligenza artificiale agentica

Di seguito sono riportati i vantaggi aggiuntivi forniti dal sistema di intelligenza artificiale agentica nello Scenario A:

- Scalabilità: può gestire picchi di volume senza costi aggiuntivi
- Disponibilità: funzionamento 24 ore su 24, 7 giorni su 7 con risposta immediata
- Coerenza: applicazione di criteri di screening uniformi
- Efficienza temporale: notevolmente ridotta time-to-hire
- Esperienza utente: feedback immediato ai candidati

Scenario B: tempo di proiezione di 15 minuti

I modelli dello scenario B hanno ottimizzato le operazioni di reclutamento in cui i reclutatori umani hanno semplificato il processo di selezione a 15 minuti per candidatura. Ciò rappresenta un miglioramento dell'efficienza del 66% rispetto allo Scenario A. Questo scenario mantiene lo stesso costo annuo a pieno carico di 112.250 dollari per un reclutatore di medio livello. Tuttavia, dimostra

un miglioramento significativo della produttività umana, con una capacità giornaliera che sale a 32 applicazioni in un turno di 8 ore e un throughput mensile che raggiunge le 660 applicazioni. La maggiore efficienza umana riduce il costo per applicazione da 45 a 15 dollari, riducendo il divario economico con il sistema di intelligenza artificiale agentica. Tuttavia, l'agente mantiene i suoi vantaggi strutturali: tempo di elaborazione di 5 minuti, disponibilità 24 ore su 24, 7 giorni su 7, che consente 288 applicazioni giornaliere, un tasso di errore del 2% inferiore rispetto al 5% umano e una capacità mensile superiore a 8.600 applicazioni. Sebbene questo miglioramento dell'efficienza prolunghi il periodo di pareggio da 1,16 mesi a 4,76 mesi e riduca i risparmi mensili da 19.833 dollari a 4.833 dollari, l'analisi rivela che i sistemi di agenti rimangono economicamente sostenibili anche quando competono con operazioni umane altamente ottimizzate: una visione fondamentale per le organizzazioni che valutano se i loro attuali livelli di efficienza dei processi giustificano gli investimenti in intelligenza artificiale agentica.

Struttura dei costi di base

La tabella seguente mostra i costi fissi annuali per lo scenario B.

Componente	Operazioni umane	Sistema di intelligenza artificiale agentica
Stipendio base	65.000 dollari	N/D
Vantaggi (30%)	19.500\$	N/D
Spazio di lavoro e attrezzature	12.000\$	N/D
Supervisione della gestione (15%)	9.750\$	N/D
Formazione e sviluppo	6.000\$	N/D
Costo fisso annuo totale	\$112.250	N/D

La tabella seguente mostra i costi di implementazione per lo scenario B.

Componente	Operazioni umane	Sistema di intelligenza artificiale agentica
------------	------------------	--

Configurazione iniziale dell'	N/D	23.000 dollari
Costi fissi mensili	9.354\$	\$500

Parametri operativi

La tabella seguente mostra le metriche operative per lo scenario B.

Parametro	Operazioni umane	Sistema di intelligenza artificiale agentica
Tempo di elaborazione per applicazione	15 minuti	5 minuti
Capacità oraria	4 applicazioni	12 applicazioni
Capacità giornaliera (turno di 8 ore)	32 applicazioni	288 applicazioni
Capacità mensile	660 applicazioni	8.640 applicazioni
Costo per applicazione	\$15	\$2,50
Costo per assunzione andata a buon fine	2.200\$	\$125
Tasso di errore	5%	2%
Costo di correzione degli errori	30 USD per errore	45 dollari per escalation

Analisi dei costi basata sul volume

La tabella seguente mostra un'analisi dei costi basata sul volume per lo scenario B. In questo esempio, il costo del sistema Agentic AI include i costi fissi e i costi di configurazione ammortizzati di 1.917 dollari al mese per 12 mesi.

Volume mensile	Costo umano	Costo del sistema Agentic AI	Risparmi mensili
100 applicazioni	1.500\$	750 dollari	750 dollari
500 applicazioni	7.500\$	2.667\$	\$4.833
1.000 applicazioni	15.000\$	4.917\$	\$10.083

Analisi del ROI

La tabella seguente mostra un'analisi del ROI per lo scenario B basata sull'elaborazione di 500 applicazioni al mese.

Parametro	Valore
Costo umano mensile	7.500 dollari
Costo mensile del sistema di intelligenza artificiale agentica	2.667\$
Risparmi mensili	4.833\$
Risparmio annuo	57.996\$
Periodo di pareggio	4,76 mesi

Confronto dei costi cumulativi

La tabella seguente mostra un confronto dei costi cumulativi per lo scenario B per i primi sei mesi, ipotizzando 500 domande al mese.

Mese	Costo umano	Costo del sistema Agentic AI	Risparmi cumulativi
1	7.500\$	\$25,667	- 18.167\$

2	\$15.000	\$28.334	- 13.334\$
3	22.500\$	\$31.001	-8.501 \$
4	30.000\$	\$33.668	-3.668 \$
5	37.500\$	\$36.335	1.165\$
6	45.000\$	\$39.002	\$5.998

Confronto tra costi e benefici per ogni scenario

Parametro	Scenario A	Scenario B	Impatto
Tempo di proiezione	45 minuti	15 minuti	Miglioramento del 66%
Capacità giornaliera	10-11 applicazioni	32 applicazioni	Aumento del 200%
Costo per applicazioni	\$45	\$15	Riduzione del 66%
Risparmio mensile (500 applicazioni)	19.833\$	\$4,833	Riduzione del 76%
Periodo di pareggio	1,16 mesi	4,76 mesi	310% più lungo

Lo scenario B dimostra significativi incrementi di efficienza nelle operazioni umane, con miglioramenti dei tempi di elaborazione che aumentano la capacità senza personale aggiuntivo e riducono sostanzialmente il costo per applicazione. Tuttavia, l'impatto finanziario rivela un quadro più sfumato: sebbene il ROI rimanga positivo, le organizzazioni devono affrontare un periodo di pareggio prolungato e risparmi mensili ridotti rispetto allo scenario A. Questi risultati evidenziano fattori decisionali critici per l'implementazione: il sistema di agenti rimane finanziariamente sostenibile anche in presenza di operazioni umane ottimizzate, ma le organizzazioni devono adottare una prospettiva di investimento a lungo termine e considerare attentamente le fluttuazioni dei volumi e le esigenze di scalabilità nella valutazione delle tempistiche di implementazione e dei rendimenti attesi.

Tuttavia, il sistema di intelligenza artificiale agentica mantiene ancora vantaggi operativi critici che vanno oltre il puro risparmio sui costi. Fornisce disponibilità 24 ore su 24, 7 giorni su 7, per un coinvolgimento immediato dei candidati indipendentemente dai fusi orari o dagli orari di lavoro. Offre una qualità di screening costante applicando criteri uniformi a ogni applicazione, ed è scalabile per gestire i picchi di volume senza incorrere in costi aggiuntivi. Offre una risposta immediata ai candidati che valorizza il marchio del datore di lavoro e l'esperienza dei candidati e opera con un fattore di affaticamento pari a zero, garantendo prestazioni di alta qualità alla prima candidatura pari alla millesima.

Gli errori umani in genere derivano da affaticamento, distrazione o lacune nelle conoscenze e spesso comportano problemi di comunicazione o informazioni errate. Gli errori del sistema Agentic AI di solito derivano da casi limite, input ambigui o limitazioni dei dati di addestramento. Questi errori tendono ad essere di natura più coerente.

Le metriche relative alla qualità e all'esperienza rivelano chiari compromessi tra le capacità umane e quelle degli agenti:

- Soddisfazione del cliente: gli esseri umani eccellono nell'empatia e nella risoluzione di problemi complessi, e gli agenti forniscono informazioni coerenti e accurate per le domande di routine.
- Tempo di risposta: i tempi di risposta favoriscono la disponibilità immediata degli agenti 24 ore su 24, 7 giorni su 7. Gli addetti forniscono assistenza durante l'orario di lavoro evitando potenziali ritardi nelle code.
- Coerenza: gli agenti forniscono risposte identiche a domande simili. Gli esseri umani possono variare nell'approccio e nell'applicazione delle conoscenze.
- Gestione dell'escalation: le questioni complesse che richiedono giudizio, creatività o intelligenza emotiva rimangono punti di forza umani.

Conclusioni e risorse

L'economia dei sistemi umani rispetto ai sistemi di intelligenza artificiale agentici rappresenta più di una decisione tecnologica. Riflette una trasformazione fondamentale nel modo in cui le organizzazioni creano valore, gestiscono il rischio e ottengono un vantaggio competitivo. Il successo richiede una valutazione sistematica delle caratteristiche del lavoro, una misurazione completa dei risultati (compresi i fattori di rischio) e una scalabilità strategica basata su risultati comprovati.

[State of Enterprise AI Adoption](#) (rapporto ISG 2025) rivela che la maggior parte delle implementazioni di intelligenza artificiale falliscono a causa di lacune di apprendimento: sistemi che non possono adattarsi, ricordare il contesto o migliorare nel tempo. Organizzazioni che raggiungono il successo si concentrano su sistemi in grado di apprendere che si integrano profondamente nei flussi di lavoro e dimostrano un miglioramento continuo attraverso il feedback umano e l'esperienza operativa.

Organizzazioni che comprendono questi principi, partendo dai lavori appropriati, scomponendo i lavori in attività, misurando tutto compreso l'impatto sul rischio e scalando ciò che funziona, otterranno un vantaggio competitivo sostenibile attraverso un utilizzo ottimale delle risorse e un'automazione incentrata sui risultati che cresce con il successo aziendale.

Il futuro appartiene alle organizzazioni in grado di combinare in modo intelligente l'esperienza umana con le funzionalità di intelligenza artificiale agentic. Questo crea modelli ibridi che offrono risultati superiori pur mantenendo la flessibilità, la capacità di apprendimento e i vantaggi collaborativi necessari per condizioni di mercato dinamiche.

Resources

Le seguenti risorse possono aiutarti a pianificare, progettare e implementare sistemi di intelligenza artificiale agentic su: AWS

- [Creazione di architetture serverless per l'IA agentic](#) su (Prescriptive Guidance) AWSAWS
- [Rendere operativa l'IA agentic](#) su (Prescriptive Guidance) AWSAWS
- [Modelli e flussi di lavoro di intelligenza artificiale agentic su \(Prescriptive Guidance\)](#) AWSAWS
- [Agentic AI](#) (guida prescrittiva)AWS
- [Centrale ottimizzazione costi AWS](#) (Servizio AWS)
- [Documentazione Amazon Bedrock](#) ()Servizio AWS
- [Pilastro dell'ottimizzazione dei costi \(AWS Well-Architected Framework\)](#)

- [Agenti e soluzioni di intelligenza artificiale](#) ()Marketplace AWS

Cronologia dei documenti

La tabella seguente descrive le modifiche significative apportate a questa guida. Per ricevere notifiche sugli aggiornamenti futuri, puoi abbonarti a un [feed RSS](#).

Modifica	Descrizione	Data
Pubblicazione iniziale	—	28 gennaio 2026

AWS Glossario delle linee guida prescrittive

I seguenti sono termini di uso comune nelle strategie, nelle guide e nei modelli forniti da AWS Prescriptive Guidance. Per suggerire voci, utilizza il link [Fornisci feedback](#) alla fine del glossario.

Numeri

7 R

Sette strategie di migrazione comuni per trasferire le applicazioni sul cloud. Queste strategie si basano sulle 5 R identificate da Gartner nel 2011 e sono le seguenti:

- **Rifattorizzare/riprogettare:** trasferisci un'applicazione e modifica la sua architettura sfruttando appieno le funzionalità native del cloud per migliorare l'agilità, le prestazioni e la scalabilità. Ciò comporta in genere la portabilità del sistema operativo e del database. Esempio: migra il tuo database Oracle locale all'edizione compatibile con Amazon Aurora PostgreSQL.
- **Ridefinire la piattaforma (lift and reshape):** trasferisci un'applicazione nel cloud e introduci un certo livello di ottimizzazione per sfruttare le funzionalità del cloud. Esempio: migra il tuo database Oracle locale ad Amazon Relational Database Service (Amazon RDS) per Oracle in Cloud AWS
- **Riacquistare (drop and shop):** passa a un prodotto diverso, in genere effettuando la transizione da una licenza tradizionale a un modello SaaS. Esempio: migra il tuo sistema di gestione delle relazioni con i clienti (CRM) su Salesforce.com.
- **Eseguire il rehosting (lift and shift):** trasferisci un'applicazione sul cloud senza apportare modifiche per sfruttare le funzionalità del cloud. Esempio: migra il tuo database Oracle locale a Oracle su un'istanza EC2 in Cloud AWS
- **Trasferire (eseguire il rehosting a livello hypervisor):** trasferisci l'infrastruttura sul cloud senza acquistare nuovo hardware, riscrivere le applicazioni o modificare le operazioni esistenti. Esegui la migrazione dei server da una piattaforma locale a un servizio cloud per la stessa piattaforma. Esempio: migra un'applicazione su Microsoft Hyper-V. AWS
- **Riesaminare (mantenere):** mantieni le applicazioni nell'ambiente di origine. Queste potrebbero includere applicazioni che richiedono una rifattorizzazione significativa che desideri rimandare a un momento successivo e applicazioni legacy che desideri mantenere, perché non vi è alcuna giustificazione aziendale per effettuarne la migrazione.
- **Ritirare:** disattiva o rimuovi le applicazioni che non sono più necessarie nell'ambiente di origine.

A

ABAC

Vedi controllo degli accessi [basato sugli attributi](#).

servizi astratti

Vedi [servizi gestiti](#).

ACIDO

Vedi [atomicità, consistenza, isolamento, durata](#).

migrazione attiva-attiva

Un metodo di migrazione del database in cui i database di origine e di destinazione vengono mantenuti sincronizzati (utilizzando uno strumento di replica bidirezionale o operazioni di doppia scrittura) ed entrambi i database gestiscono le transazioni provenienti dalle applicazioni di connessione durante la migrazione. Questo metodo supporta la migrazione in piccoli batch controllati anziché richiedere una conversione una tantum. È più flessibile ma richiede più lavoro rispetto alla migrazione [attiva-passiva](#).

migrazione attiva-passiva

Un metodo di migrazione del database in cui i database di origine e di destinazione vengono mantenuti sincronizzati, ma solo il database di origine gestisce le transazioni provenienti dalle applicazioni di connessione mentre i dati vengono replicati nel database di destinazione. Il database di destinazione non accetta alcuna transazione durante la migrazione.

funzione di aggregazione

Una funzione SQL che opera su un gruppo di righe e calcola un singolo valore restituito per il gruppo. Esempi di funzioni aggregate includono SUM e MAX.

Intelligenza artificiale

Vedi [intelligenza artificiale](#).

AIOps

Guarda le [operazioni di intelligenza artificiale](#).

anonimizzazione

Il processo di eliminazione permanente delle informazioni personali in un set di dati.

L'anonimizzazione può aiutare a proteggere la privacy personale. I dati anonimi non sono più considerati dati personali.

anti-modello

Una soluzione utilizzata frequentemente per un problema ricorrente in cui la soluzione è controproducente, inefficace o meno efficace di un'alternativa.

controllo delle applicazioni

Un approccio alla sicurezza che consente l'uso solo di applicazioni approvate per proteggere un sistema dal malware.

portfolio di applicazioni

Una raccolta di informazioni dettagliate su ogni applicazione utilizzata da un'organizzazione, compresi i costi di creazione e manutenzione dell'applicazione e il relativo valore aziendale. Queste informazioni sono fondamentali per [il processo di scoperta e analisi del portfolio](#) e aiutano a identificare e ad assegnare la priorità alle applicazioni da migrare, modernizzare e ottimizzare.

intelligenza artificiale (IA)

Il campo dell'informatica dedicato all'uso delle tecnologie informatiche per svolgere funzioni cognitive tipicamente associate agli esseri umani, come l'apprendimento, la risoluzione di problemi e il riconoscimento di schemi. Per ulteriori informazioni, consulta la sezione [Che cos'è l'intelligenza artificiale?](#)

operazioni di intelligenza artificiale (AIOps)

Il processo di utilizzo delle tecniche di machine learning per risolvere problemi operativi, ridurre gli incidenti operativi e l'intervento umano e aumentare la qualità del servizio. Per ulteriori informazioni su come AIOps viene utilizzata nella strategia di AWS migrazione, consulta la [guida all'integrazione delle operazioni](#).

crittografia asimmetrica

Un algoritmo di crittografia che utilizza una coppia di chiavi, una chiave pubblica per la crittografia e una chiave privata per la decrittografia. Puoi condividere la chiave pubblica perché non viene utilizzata per la decrittografia, ma l'accesso alla chiave privata deve essere altamente limitato.

atomicità, consistenza, isolamento, durabilità (ACID)

Un insieme di proprietà del software che garantiscono la validità dei dati e l'affidabilità operativa di un database, anche in caso di errori, interruzioni di corrente o altri problemi.

Controllo degli accessi basato su attributi (ABAC)

La pratica di creare autorizzazioni dettagliate basate su attributi utente, come reparto, ruolo professionale e nome del team. Per ulteriori informazioni, consulta [ABAC AWS](#) nella documentazione AWS Identity and Access Management (IAM).

fonte di dati autorevole

Una posizione in cui è archiviata la versione principale dei dati, considerata la fonte di informazioni più affidabile. È possibile copiare i dati dalla fonte di dati autorevole in altre posizioni allo scopo di elaborarli o modificarli, ad esempio anonimizzandoli, oscurandoli o pseudonimizzandoli.

Zona di disponibilità

Una posizione distinta all'interno di un edificio Regione AWS che è isolata dai guasti in altre zone di disponibilità e offre una connettività di rete economica e a bassa latenza verso altre zone di disponibilità nella stessa regione.

AWS Cloud Adoption Framework (CAF)AWS

Un framework di linee guida e best practice AWS per aiutare le organizzazioni a sviluppare un piano efficiente ed efficace per passare con successo al cloud. AWS CAF organizza le linee guida in sei aree di interesse chiamate prospettive: business, persone, governance, piattaforma, sicurezza e operazioni. Le prospettive relative ad azienda, persone e governance si concentrano sulle competenze e sui processi aziendali; le prospettive relative alla piattaforma, alla sicurezza e alle operazioni si concentrano sulle competenze e sui processi tecnici. Ad esempio, la prospettiva relativa alle persone si rivolge alle parti interessate che gestiscono le risorse umane (HR), le funzioni del personale e la gestione del personale. In questa prospettiva, AWS CAF fornisce linee guida per lo sviluppo delle persone, la formazione e le comunicazioni per aiutare a preparare l'organizzazione all'adozione del cloud di successo. Per ulteriori informazioni, consulta il [sito web di AWS CAF](#) e il [white paper AWS CAF](#).

AWS Workload Qualification Framework (WQF)AWS

Uno strumento che valuta i carichi di lavoro di migrazione dei database, consiglia strategie di migrazione e fornisce stime del lavoro. AWS WQF è incluso in (). AWS Schema Conversion Tool AWS SCT Analizza gli schemi di database e gli oggetti di codice, il codice dell'applicazione, le dipendenze e le caratteristiche delle prestazioni e fornisce report di valutazione.

B

bot difettoso

Un [bot](#) che ha lo scopo di interrompere o causare danni a individui o organizzazioni.

BCP

Vedi la [pianificazione della continuità operativa](#).

grafico comportamentale

Una vista unificata, interattiva dei comportamenti delle risorse e delle interazioni nel tempo. Puoi utilizzare un grafico comportamentale con Amazon Detective per esaminare tentativi di accesso non riusciti, chiamate API sospette e azioni simili. Per ulteriori informazioni, consulta [Dati in un grafico comportamentale](#) nella documentazione di Detective.

sistema big-endian

Un sistema che memorizza per primo il byte più importante. Vedi anche [endianness](#).

Classificazione binaria

Un processo che prevede un risultato binario (una delle due classi possibili). Ad esempio, il modello di machine learning potrebbe dover prevedere problemi come "Questa e-mail è spam o non è spam?" o "Questo prodotto è un libro o un'auto?"

filtro Bloom

Una struttura di dati probabilistica ed efficiente in termini di memoria che viene utilizzata per verificare se un elemento fa parte di un set.

implementazione blu/verde

Una strategia di implementazione in cui si creano due ambienti separati ma identici. La versione corrente dell'applicazione viene eseguita in un ambiente (blu) e la nuova versione dell'applicazione nell'altro ambiente (verde). Questa strategia consente di ripristinare rapidamente il sistema con un impatto minimo.

bot

Un'applicazione software che esegue attività automatizzate su Internet e simula l'attività o l'interazione umana. Alcuni bot sono utili o utili, come i web crawler che indicizzano le informazioni su Internet. Alcuni altri bot, noti come bot dannosi, hanno lo scopo di disturbare o causare danni a individui o organizzazioni.

botnet

Reti di [bot](#) infettate da [malware](#) e controllate da un'unica parte, nota come bot herder o bot operator. Le botnet sono il meccanismo più noto per scalare i bot e il loro impatto.

ramo

Un'area contenuta di un repository di codice. Il primo ramo creato in un repository è il ramo principale. È possibile creare un nuovo ramo a partire da un ramo esistente e quindi sviluppare funzionalità o correggere bug al suo interno. Un ramo creato per sviluppare una funzionalità viene comunemente detto ramo di funzionalità. Quando la funzionalità è pronta per il rilascio, il ramo di funzionalità viene ricongiunto al ramo principale. Per ulteriori informazioni, consulta [Informazioni sulle filiali](#) (documentazione). GitHub

accesso break-glass

In circostanze eccezionali e tramite una procedura approvata, un mezzo rapido per consentire a un utente di accedere a un sito a Account AWS cui in genere non dispone delle autorizzazioni necessarie. Per ulteriori informazioni, vedere l'indicatore [Implementate break-glass procedures](#) nella guida Well-Architected AWS .

strategia brownfield

L'infrastruttura esistente nell'ambiente. Quando si adotta una strategia brownfield per un'architettura di sistema, si progetta l'architettura in base ai vincoli dei sistemi e dell'infrastruttura attuali. Per l'espansione dell'infrastruttura esistente, è possibile combinare strategie brownfield e [greenfield](#).

cache del buffer

L'area di memoria in cui sono archiviati i dati a cui si accede con maggiore frequenza.

capacità di business

Azioni intraprese da un'azienda per generare valore (ad esempio vendite, assistenza clienti o marketing). Le architetture dei microservizi e le decisioni di sviluppo possono essere guidate dalle capacità aziendali. Per ulteriori informazioni, consulta la sezione [Organizzazione in base alle funzionalità aziendali](#) del whitepaper [Esecuzione di microservizi containerizzati su AWS](#).

pianificazione della continuità operativa (BCP)

Un piano che affronta il potenziale impatto di un evento che comporta l'interruzione dell'attività, come una migrazione su larga scala, sulle operazioni e consente a un'azienda di riprendere rapidamente le operazioni.

C

CAF

Vedi [Cloud Adoption AWS Framework](#).

implementazione canaria

Il rilascio lento e incrementale di una versione agli utenti finali. Quando sei sicuro, distribuisce la nuova versione e sostituisci la versione corrente nella sua interezza.

CCoE

Vedi [Cloud Center of Excellence](#).

CDC

Vedi [Change Data Capture](#).

Change Data Capture (CDC)

Il processo di tracciamento delle modifiche a un'origine dati, ad esempio una tabella di database, e di registrazione dei metadati relativi alla modifica. È possibile utilizzare CDC per vari scopi, ad esempio il controllo o la replica delle modifiche in un sistema di destinazione per mantenere la sincronizzazione.

ingegneria del caos

Introduzione intenzionale di guasti o eventi dirompenti per testare la resilienza di un sistema. Puoi usare [AWS Fault Injection Service \(AWS FIS\)](#) per eseguire esperimenti che stressano i tuoi AWS carichi di lavoro e valutarne la risposta.

CI/CD

Vedi [integrazione continua e distribuzione continua](#).

classificazione

Un processo di categorizzazione che aiuta a generare previsioni. I modelli di ML per problemi di classificazione prevedono un valore discreto. I valori discreti sono sempre distinti l'uno dall'altro. Ad esempio, un modello potrebbe dover valutare se in un'immagine è presente o meno un'auto.

crittografia lato client

Crittografia dei dati a livello locale, prima che il destinatario li Servizio AWS riceva.

Centro di eccellenza cloud (CCoE)

Un team multidisciplinare che guida le iniziative di adozione del cloud in tutta l'organizzazione, tra cui lo sviluppo di best practice per il cloud, la mobilitazione delle risorse, la definizione delle tempistiche di migrazione e la guida dell'organizzazione attraverso trasformazioni su larga scala. Per ulteriori informazioni, consulta gli [CCoE post](#) sull' Cloud AWS Enterprise Strategy Blog.

cloud computing

La tecnologia cloud generalmente utilizzata per l'archiviazione remota di dati e la gestione dei dispositivi IoT. Il cloud computing è generalmente collegato alla tecnologia di [edge computing](#).

modello operativo cloud

In un'organizzazione IT, il modello operativo utilizzato per creare, maturare e ottimizzare uno o più ambienti cloud. Per ulteriori informazioni, consulta [Building your Cloud Operating Model](#).

fasi di adozione del cloud

Le quattro fasi che le organizzazioni in genere attraversano quando migrano verso Cloud AWS:

- Progetto: esecuzione di alcuni progetti relativi al cloud per scopi di dimostrazione e apprendimento
- Fondamento: effettuare investimenti fondamentali per scalare l'adozione del cloud (ad esempio, creazione di una landing zone, definizione di una CCo E, definizione di un modello operativo)
- Migrazione: migrazione di singole applicazioni
- Reinvenzione: ottimizzazione di prodotti e servizi e innovazione nel cloud

Queste fasi sono state definite da Stephen Orban nel post sul blog The [Journey Toward Cloud-First & the Stages of Adoption on the Enterprise Strategy](#). Cloud AWS [Per informazioni su come si relazionano alla strategia di AWS migrazione, consulta la guida alla preparazione alla migrazione.](#)

CMDB

Vedi [database di gestione della configurazione](#).

repository di codice

Una posizione in cui il codice di origine e altri asset, come documentazione, esempi e script, vengono archiviati e aggiornati attraverso processi di controllo delle versioni. Gli archivi cloud più comuni includono GitHub oBitbucket Cloud. Ogni versione del codice è denominata ramo. In una struttura a microservizi, ogni repository è dedicato a una singola funzionalità. Una singola pipeline CI/CD può utilizzare più repository.

cache fredda

Una cache del buffer vuota, non ben popolata o contenente dati obsoleti o irrilevanti. Ciò influisce sulle prestazioni perché l'istanza di database deve leggere dalla memoria o dal disco principale, il che richiede più tempo rispetto alla lettura dalla cache del buffer.

dati freddi

Dati a cui si accede raramente e che in genere sono storici. Quando si eseguono interrogazioni di questo tipo di dati, le interrogazioni lente sono in genere accettabili. Lo spostamento di questi dati su livelli o classi di storage meno costosi e con prestazioni inferiori può ridurre i costi.

visione artificiale (CV)

Un campo dell'[intelligenza artificiale](#) che utilizza l'apprendimento automatico per analizzare ed estrarre informazioni da formati visivi come immagini e video digitali. Ad esempio, Amazon SageMaker AI fornisce algoritmi di elaborazione delle immagini per CV.

deriva della configurazione

Per un carico di lavoro, una modifica della configurazione rispetto allo stato previsto. Potrebbe causare la non conformità del carico di lavoro e in genere è graduale e involontaria.

database di gestione della configurazione (CMDB)

Un repository che archivia e gestisce le informazioni su un database e il relativo ambiente IT, inclusi i componenti hardware e software e le relative configurazioni. In genere si utilizzano i dati di un CMDB nella fase di individuazione e analisi del portafoglio della migrazione.

Pacchetto di conformità

Una raccolta di AWS Config regole e azioni correttive che puoi assemblare per personalizzare i controlli di conformità e sicurezza. È possibile distribuire un pacchetto di conformità come singola entità in una regione Account AWS and o all'interno di un'organizzazione utilizzando un modello YAML. Per ulteriori informazioni, consulta i [Conformance](#) pack nella documentazione. AWS Config

integrazione e distribuzione continua (continuous integration and continuous delivery, CI/CD)

Il processo di automazione delle fasi di origine, compilazione, test, gestione temporanea e produzione del processo di rilascio del software. CI/CD viene comunemente descritto come una pipeline. CI/CD può aiutarvi ad automatizzare i processi, migliorare la produttività, migliorare la qualità del codice e velocizzare le consegne. Per ulteriori informazioni, consulta [Vantaggi della distribuzione continua](#). CD può anche significare continuous deployment (implementazione continua). Per ulteriori informazioni, consulta [Distribuzione continua e implementazione continua a confronto](#).

CV

Vedi [visione artificiale](#).

D

dati a riposo

Dati stazionari nella rete, ad esempio i dati archiviati.

classificazione dei dati

Un processo per identificare e classificare i dati nella rete in base alla loro criticità e sensibilità. È un componente fondamentale di qualsiasi strategia di gestione dei rischi di sicurezza informatica perché consente di determinare i controlli di protezione e conservazione appropriati per i dati. La classificazione dei dati è un componente del pilastro della sicurezza nel AWS Well-Architected Framework. Per ulteriori informazioni, consulta [Classificazione dei dati](#).

deriva dei dati

Una variazione significativa tra i dati di produzione e i dati utilizzati per addestrare un modello di machine learning o una modifica significativa dei dati di input nel tempo. La deriva dei dati può ridurre la qualità, l'accuratezza e l'equità complessive nelle previsioni dei modelli ML.

dati in transito

Dati che si spostano attivamente attraverso la rete, ad esempio tra le risorse di rete.

rete di dati

Un framework architettonico che fornisce la proprietà distribuita e decentralizzata dei dati con gestione e governance centralizzate.

riduzione al minimo dei dati

Il principio della raccolta e del trattamento dei soli dati strettamente necessari. Praticare la riduzione al minimo dei dati in the Cloud AWS può ridurre i rischi per la privacy, i costi e l'impronta di carbonio delle analisi.

perimetro dei dati

Una serie di barriere preventive nell' AWS ambiente che aiutano a garantire che solo le identità attendibili accedano alle risorse attendibili delle reti previste. Per ulteriori informazioni, consulta [Building a data perimeter](#) on AWS.

pre-elaborazione dei dati

Trasformare i dati grezzi in un formato che possa essere facilmente analizzato dal modello di ML. La pre-elaborazione dei dati può comportare la rimozione di determinate colonne o righe e l'eliminazione di valori mancanti, incoerenti o duplicati.

provenienza dei dati

Il processo di tracciamento dell'origine e della cronologia dei dati durante il loro ciclo di vita, ad esempio il modo in cui i dati sono stati generati, trasmessi e archiviati.

soggetto dei dati

Un individuo i cui dati vengono raccolti ed elaborati.

data warehouse

Un sistema di gestione dei dati che supporta la business intelligence, come l'analisi. I data warehouse contengono in genere grandi quantità di dati storici e vengono generalmente utilizzati per interrogazioni e analisi.

linguaggio di definizione del database (DDL)

Istruzioni o comandi per creare o modificare la struttura di tabelle e oggetti in un database.

linguaggio di manipolazione del database (DML)

Istruzioni o comandi per modificare (inserire, aggiornare ed eliminare) informazioni in un database.

DDL

Vedi linguaggio di [definizione del database](#).

deep ensemble

Combinare più modelli di deep learning per la previsione. È possibile utilizzare i deep ensemble per ottenere una previsione più accurata o per stimare l'incertezza nelle previsioni.

deep learning

Un sottocampo del ML che utilizza più livelli di reti neurali artificiali per identificare la mappatura tra i dati di input e le variabili target di interesse.

defense-in-depth

Un approccio alla sicurezza delle informazioni in cui una serie di meccanismi e controlli di sicurezza sono accuratamente stratificati su una rete di computer per proteggere la riservatezza,

l'integrità e la disponibilità della rete e dei dati al suo interno. Quando si adotta questa strategia AWS, si aggiungono più controlli a diversi livelli della AWS Organizations struttura per proteggere le risorse. Ad esempio, un defense-in-depth approccio potrebbe combinare l'autenticazione a più fattori, la segmentazione della rete e la crittografia.

amministratore delegato

In AWS Organizations, un servizio compatibile può registrare un account AWS membro per amministrare gli account dell'organizzazione e gestire le autorizzazioni per quel servizio. Questo account è denominato amministratore delegato per quel servizio specifico. Per ulteriori informazioni e un elenco di servizi compatibili, consulta [Servizi che funzionano con AWS Organizations](#) nella documentazione di AWS Organizations .

implementazione

Il processo di creazione di un'applicazione, di nuove funzionalità o di correzioni di codice disponibili nell'ambiente di destinazione. L'implementazione prevede l'applicazione di modifiche in una base di codice, seguita dalla creazione e dall'esecuzione di tale base di codice negli ambienti applicativi.

Ambiente di sviluppo

[Vedi ambiente.](#)

controllo di rilevamento

Un controllo di sicurezza progettato per rilevare, registrare e avvisare dopo che si è verificato un evento. Questi controlli rappresentano una seconda linea di difesa e avvisano l'utente in caso di eventi di sicurezza che aggirano i controlli preventivi in vigore. Per ulteriori informazioni, consulta [Controlli di rilevamento](#) in Implementazione dei controlli di sicurezza in AWS.

mappatura del flusso di valore dello sviluppo (DVSM)

Un processo utilizzato per identificare e dare priorità ai vincoli che influiscono negativamente sulla velocità e sulla qualità nel ciclo di vita dello sviluppo del software. DVSM estende il processo di mappatura del flusso di valore originariamente progettato per pratiche di produzione snella. Si concentra sulle fasi e sui team necessari per creare e trasferire valore attraverso il processo di sviluppo del software.

gemello digitale

Una rappresentazione virtuale di un sistema reale, ad esempio un edificio, una fabbrica, un'attrezzatura industriale o una linea di produzione. I gemelli digitali supportano la manutenzione predittiva, il monitoraggio remoto e l'ottimizzazione della produzione.

tabella delle dimensioni

In uno [schema a stella](#), una tabella più piccola che contiene gli attributi dei dati quantitativi in una tabella dei fatti. Gli attributi della tabella delle dimensioni sono in genere campi di testo o numeri discreti che si comportano come testo. Questi attributi vengono comunemente utilizzati per il vincolo delle query, il filtraggio e l'etichettatura dei set di risultati.

disastro

Un evento che impedisce a un carico di lavoro o a un sistema di raggiungere gli obiettivi aziendali nella sua sede principale di implementazione. Questi eventi possono essere disastri naturali, guasti tecnici o il risultato di azioni umane, come errori di configurazione involontari o attacchi di malware.

disaster recovery (DR)

La strategia e il processo utilizzati per ridurre al minimo i tempi di inattività e la perdita di dati causati da un [disastro](#). Per ulteriori informazioni, consulta [Disaster Recovery of Workloads su AWS: Recovery in the Cloud in the AWS Well-Architected Framework](#).

DML

Vedi linguaggio di manipolazione [del database](#).

progettazione basata sul dominio

Un approccio allo sviluppo di un sistema software complesso collegandone i componenti a domini in evoluzione, o obiettivi aziendali principali, perseguiti da ciascun componente. Questo concetto è stato introdotto da Eric Evans nel suo libro, *Domain-Driven Design: Tackling Complexity in the Heart of Software* (Boston: Addison-Wesley Professional, 2003). Per informazioni su come utilizzare la progettazione basata sul dominio con il modello del fico strangolatore (Strangler Fig), consulta la sezione [Modernizzazione incrementale dei servizi Web Microsoft ASP.NET \(ASMX\) legacy utilizzando container e il Gateway Amazon API](#).

DOTT.

Vedi [disaster recovery](#).

rilevamento della deriva

Tracciamento delle deviazioni da una configurazione di base. Ad esempio, è possibile AWS CloudFormation utilizzarlo per [rilevare deviazioni nelle risorse di sistema](#) oppure AWS Control Tower per [rilevare cambiamenti nella landing zone](#) che potrebbero influire sulla conformità ai requisiti di governance.

DVSM

Vedi la [mappatura del flusso di valore dello sviluppo](#).

E

EDA

Vedi [analisi esplorativa dei dati](#).

MODIFICA

Vedi [scambio elettronico di dati](#).

edge computing

La tecnologia che aumenta la potenza di calcolo per i dispositivi intelligenti all'edge di una rete IoT. Rispetto al [cloud computing](#), [l'edge computing](#) può ridurre la latenza di comunicazione e migliorare i tempi di risposta.

scambio elettronico di dati (EDI)

Lo scambio automatizzato di documenti aziendali tra organizzazioni. Per ulteriori informazioni, vedere [Cos'è lo scambio elettronico di dati](#).

crittografia

Un processo di elaborazione che trasforma i dati in chiaro, leggibili dall'uomo, in testo cifrato.

chiave crittografica

Una stringa crittografica di bit randomizzati generata da un algoritmo di crittografia. Le chiavi possono variare di lunghezza e ogni chiave è progettata per essere imprevedibile e univoca.

endianità

L'ordine in cui i byte vengono archiviati nella memoria del computer. I sistemi big-endian memorizzano per primo il byte più importante. I sistemi little-endian memorizzano per primo il byte meno importante.

endpoint

Vedi [service endpoint](#).

servizio endpoint

Un servizio che puoi ospitare in un cloud privato virtuale (VPC) da condividere con altri utenti. Puoi creare un servizio endpoint con AWS PrivateLink e concedere autorizzazioni ad altri Account AWS o a AWS Identity and Access Management (IAM) principali. Questi account o principali possono connettersi al servizio endpoint in privato creando endpoint VPC di interfaccia. Per ulteriori informazioni, consulta [Creazione di un servizio endpoint](#) nella documentazione di Amazon Virtual Private Cloud (Amazon VPC).

pianificazione delle risorse aziendali (ERP)

Un sistema che automatizza e gestisce i processi aziendali chiave (come contabilità, [MES](#) e gestione dei progetti) per un'azienda.

crittografia envelope

Il processo di crittografia di una chiave di crittografia con un'altra chiave di crittografia. Per ulteriori informazioni, vedete [Envelope encryption](#) nella documentazione AWS Key Management Service (AWS KMS).

ambiente

Un'istanza di un'applicazione in esecuzione. Di seguito sono riportati i tipi di ambiente più comuni nel cloud computing:

- ambiente di sviluppo: un'istanza di un'applicazione in esecuzione disponibile solo per il team principale responsabile della manutenzione dell'applicazione. Gli ambienti di sviluppo vengono utilizzati per testare le modifiche prima di promuoverle negli ambienti superiori. Questo tipo di ambiente viene talvolta definito ambiente di test.
- ambienti inferiori: tutti gli ambienti di sviluppo di un'applicazione, ad esempio quelli utilizzati per le build e i test iniziali.
- ambiente di produzione: un'istanza di un'applicazione in esecuzione a cui gli utenti finali possono accedere. In una CI/CD pipeline, l'ambiente di produzione è l'ultimo ambiente di distribuzione.
- ambienti superiori: tutti gli ambienti a cui possono accedere utenti diversi dal team di sviluppo principale. Si può trattare di un ambiente di produzione, ambienti di riproduzione e ambienti per i test di accettazione da parte degli utenti.

epica

Nelle metodologie agili, categorie funzionali che aiutano a organizzare e dare priorità al lavoro. Le epiche forniscono una descrizione di alto livello dei requisiti e delle attività di implementazione.

Ad esempio, le epopee della sicurezza AWS CAF includono la gestione delle identità e degli accessi, i controlli investigativi, la sicurezza dell'infrastruttura, la protezione dei dati e la risposta agli incidenti. Per ulteriori informazioni sulle epiche, consulta la strategia di migrazione AWS , consulta la [guida all'implementazione del programma](#).

ERP

Vedi [pianificazione delle risorse aziendali](#).

analisi esplorativa dei dati (EDA)

Il processo di analisi di un set di dati per comprenderne le caratteristiche principali. Si raccolgono o si aggregano dati e quindi si eseguono indagini iniziali per trovare modelli, rilevare anomalie e verificare ipotesi. L'EDA viene eseguita calcolando statistiche di riepilogo e creando visualizzazioni di dati.

F

tabella dei fatti

Il tavolo centrale in uno [schema a stella](#). Memorizza dati quantitativi sulle operazioni aziendali. In genere, una tabella dei fatti contiene due tipi di colonne: quelle che contengono misure e quelle che contengono una chiave esterna per una tabella di dimensioni.

fallire velocemente

Una filosofia che utilizza test frequenti e incrementali per ridurre il ciclo di vita dello sviluppo. È una parte fondamentale di un approccio agile.

limite di isolamento dei guasti

Nel Cloud AWS, un limite come una zona di disponibilità Regione AWS, un piano di controllo o un piano dati che limita l'effetto di un errore e aiuta a migliorare la resilienza dei carichi di lavoro. Per ulteriori informazioni, consulta [AWS Fault Isolation Boundaries](#).

ramo di funzionalità

Vedi [filiale](#).

caratteristiche

I dati di input che usi per fare una previsione. Ad esempio, in un contesto di produzione, le caratteristiche potrebbero essere immagini acquisite periodicamente dalla linea di produzione.

importanza delle caratteristiche

Quanto è importante una caratteristica per le previsioni di un modello. Di solito viene espresso come punteggio numerico che può essere calcolato con varie tecniche, come Shapley Additive Explanations (SHAP) e gradienti integrati. Per ulteriori informazioni, consulta [Interpretabilità del modello di machine learning con AWS](#).

trasformazione delle funzionalità

Per ottimizzare i dati per il processo di machine learning, incluso l'arricchimento dei dati con fonti aggiuntive, il dimensionamento dei valori o l'estrazione di più set di informazioni da un singolo campo di dati. Ciò consente al modello di ML di trarre vantaggio dai dati. Ad esempio, se suddividi la data "2021-05-27 00:15:37" in "2021", "maggio", "giovedì" e "15", puoi aiutare l'algoritmo di apprendimento ad apprendere modelli sfumati associati a diversi componenti dei dati.

prompt con pochi scatti

Fornire a un [LLM](#) un numero limitato di esempi che dimostrino l'attività e il risultato desiderato prima di chiedergli di eseguire un'attività simile. Questa tecnica è un'applicazione dell'apprendimento contestuale, in cui i modelli imparano da esempi (immagini) incorporati nei prompt. I prompt con pochi passaggi possono essere efficaci per attività che richiedono una formattazione, un ragionamento o una conoscenza del dominio specifici. [Vedi anche zero-shot prompting](#).

FGAC

Vedi il controllo [granulare degli accessi](#).

controllo granulare degli accessi (FGAC)

L'uso di più condizioni per consentire o rifiutare una richiesta di accesso.

migrazione flash-cut

Un metodo di migrazione del database che utilizza la replica continua dei dati tramite [l'acquisizione dei dati delle modifiche](#) per migrare i dati nel più breve tempo possibile, anziché utilizzare un approccio graduale. L'obiettivo è ridurre al minimo i tempi di inattività.

FM

[Vedi modello di base](#).

modello di fondazione (FM)

Una grande rete neurale di deep learning che si è addestrata su enormi set di dati generalizzati e non etichettati. FMs sono in grado di svolgere un'ampia varietà di attività generali, come

comprendere il linguaggio, generare testo e immagini e conversare in linguaggio naturale. Per ulteriori informazioni, consulta [Cosa sono i modelli Foundation](#).

G

IA generativa

Un sottoinsieme di modelli di [intelligenza artificiale](#) che sono stati addestrati su grandi quantità di dati e che possono utilizzare un semplice messaggio di testo per creare nuovi contenuti e artefatti, come immagini, video, testo e audio. Per ulteriori informazioni, consulta [Cos'è l'IA generativa](#).

blocco geografico

Vedi [restrizioni geografiche](#).

limitazioni geografiche (blocco geografico)

In Amazon CloudFront, un'opzione per impedire agli utenti di determinati paesi di accedere alle distribuzioni di contenuti. Puoi utilizzare un elenco consentito o un elenco di blocco per specificare i paesi approvati e vietati. Per ulteriori informazioni, consulta [Limitare la distribuzione geografica dei contenuti](#) nella CloudFront documentazione.

Flusso di lavoro di GitFlow

Un approccio in cui gli ambienti inferiori e superiori utilizzano rami diversi in un repository di codice di origine. Il flusso di lavoro Gitflow è considerato obsoleto e il flusso di lavoro [basato su trunk è l'approccio moderno e preferito](#).

immagine dorata

Un'istantanea di un sistema o di un software utilizzata come modello per distribuire nuove istanze di quel sistema o software. Ad esempio, nella produzione, un'immagine dorata può essere utilizzata per fornire software su più dispositivi e contribuire a migliorare la velocità, la scalabilità e la produttività nelle operazioni di produzione dei dispositivi.

strategia greenfield

L'assenza di infrastrutture esistenti in un nuovo ambiente. Quando si adotta una strategia greenfield per un'architettura di sistema, è possibile selezionare tutte le nuove tecnologie senza il vincolo della compatibilità con l'infrastruttura esistente, nota anche come [brownfield](#). Per l'espansione dell'infrastruttura esistente, è possibile combinare strategie brownfield e greenfield.

guardrail

Una regola di alto livello che aiuta a governare le risorse, le politiche e la conformità tra le unità organizzative (). OUs I guardrail preventivi applicano le policy per garantire l'allineamento agli standard di conformità. Vengono implementati utilizzando le policy di controllo dei servizi e i limiti delle autorizzazioni IAM. I guardrail di rilevamento rilevano le violazioni delle policy e i problemi di conformità e generano avvisi per porvi rimedio. Sono implementati utilizzando Amazon AWS Config AWS Security Hub CSPM GuardDuty AWS Trusted Advisor, Amazon Inspector e controlli personalizzati AWS Lambda .

H

AH

Vedi [disponibilità elevata](#).

migrazione di database eterogenea

Migrazione del database di origine in un database di destinazione che utilizza un motore di database diverso (ad esempio, da Oracle ad Amazon Aurora). La migrazione eterogenea fa in genere parte di uno sforzo di riprogettazione e la conversione dello schema può essere un'attività complessa. [AWS offre AWS SCT](#) che aiuta con le conversioni dello schema.

alta disponibilità (HA)

La capacità di un carico di lavoro di funzionare in modo continuo, senza intervento, in caso di sfide o disastri. I sistemi HA sono progettati per il failover automatico, fornire costantemente prestazioni di alta qualità e gestire carichi e guasti diversi con un impatto minimo sulle prestazioni.

modernizzazione storica

Un approccio utilizzato per modernizzare e aggiornare i sistemi di tecnologia operativa (OT) per soddisfare meglio le esigenze dell'industria manifatturiera. Uno storico è un tipo di database utilizzato per raccogliere e archiviare dati da varie fonti in una fabbrica.

dati di blocco

[Una parte di dati storici etichettati che viene trattenuta da un set di dati utilizzata per addestrare un modello di apprendimento automatico.](#) È possibile utilizzare i dati di holdout per valutare le prestazioni del modello confrontando le previsioni del modello con i dati di holdout.

migrazione di database omogenea

Migrazione del database di origine in un database di destinazione che condivide lo stesso motore di database (ad esempio, da Microsoft SQL Server ad Amazon RDS per SQL Server). La migrazione omogenea fa in genere parte di un'operazione di rehosting o ridefinizione della piattaforma. Per migrare lo schema è possibile utilizzare le utilità native del database.

dati caldi

Dati a cui si accede frequentemente, come dati in tempo reale o dati di traduzione recenti. Questi dati richiedono in genere un livello o una classe di storage ad alte prestazioni per fornire risposte rapide alle query.

hotfix

Una soluzione urgente per un problema critico in un ambiente di produzione. A causa della sua urgenza, un hotfix viene in genere creato al di fuori del tipico DevOps flusso di lavoro di rilascio.

periodo di hypercare

Subito dopo la conversione, il periodo di tempo in cui un team di migrazione gestisce e monitora le applicazioni migrate nel cloud per risolvere eventuali problemi. In genere, questo periodo dura da 1 a 4 giorni. Al termine del periodo di hypercare, il team addetto alla migrazione in genere trasferisce la responsabilità delle applicazioni al team addetto alle operazioni cloud.

I

IaC

Vedi [l'infrastruttura come codice](#).

Policy basata su identità

Una policy associata a uno o più principi IAM che definisce le relative autorizzazioni all'interno dell'Cloud AWS ambiente.

applicazione inattiva

Un'applicazione che prevede un uso di CPU e memoria medio compreso tra il 5% e il 20% in un periodo di 90 giorni. In un progetto di migrazione, è normale ritirare queste applicazioni o mantenerle on-premise.

IIoT

Vedi [Industrial Internet of Things](#).

infrastruttura immutabile

Un modello che implementa una nuova infrastruttura per i carichi di lavoro di produzione anziché aggiornare, applicare patch o modificare l'infrastruttura esistente. [Le infrastrutture immutabili sono intrinsecamente più coerenti, affidabili e prevedibili delle infrastrutture mutabili](#). Per ulteriori informazioni, consulta la best practice [Deploy using immutable infrastructure in Well-Architected AWS Framework](#).

VPC in ingresso (ingress)

In un'architettura AWS multi-account, un VPC che accetta, ispeziona e indirizza le connessioni di rete dall'esterno di un'applicazione. La [AWS Security Reference Architecture](#) consiglia di configurare l'account di rete con funzionalità in entrata, in uscita e di ispezione VPCs per proteggere l'interfaccia bidirezionale tra l'applicazione e Internet in generale.

migrazione incrementale

Una strategia di conversione in cui si esegue la migrazione dell'applicazione in piccole parti anziché eseguire una conversione singola e completa. Ad esempio, inizialmente potresti spostare solo alcuni microservizi o utenti nel nuovo sistema. Dopo aver verificato che tutto funzioni correttamente, puoi spostare in modo incrementale microservizi o utenti aggiuntivi fino alla disattivazione del sistema legacy. Questa strategia riduce i rischi associati alle migrazioni di grandi dimensioni.

Industria 4.0

Un termine introdotto da [Klaus Schwab](#) nel 2016 per riferirsi alla modernizzazione dei processi di produzione attraverso progressi in termini di connettività, dati in tempo reale, automazione, analisi e AI/ML.

infrastruttura

Tutte le risorse e gli asset contenuti nell'ambiente di un'applicazione.

infrastruttura come codice (IaC)

Il processo di provisioning e gestione dell'infrastruttura di un'applicazione tramite un insieme di file di configurazione. Il processo IaC è progettato per aiutarti a centralizzare la gestione dell'infrastruttura, a standardizzare le risorse e a dimensionare rapidamente, in modo che i nuovi ambienti siano ripetibili, affidabili e coerenti.

IIo/Internet delle cose industriale (T)

L'uso di sensori e dispositivi connessi a Internet nei settori industriali, come quello manifatturiero, energetico, automobilistico, sanitario, delle scienze della vita e dell'agricoltura. Per ulteriori

informazioni, vedere [Creazione di una strategia di trasformazione digitale per l'Internet of Things \(IIoT\) industriale](#).

VPC di ispezione

In un'architettura AWS multi-account, un VPC centralizzato che gestisce le ispezioni del traffico di rete tra VPCs (nello stesso o in modo diverso Regioni AWS), Internet e le reti locali. La [AWS Security Reference Architecture](#) consiglia di configurare l'account di rete con informazioni in entrata, in uscita e di ispezione VPCs per proteggere l'interfaccia bidirezionale tra l'applicazione e Internet in generale.

Internet of Things (IoT)

La rete di oggetti fisici connessi con sensori o processori incorporati che comunicano con altri dispositivi e sistemi tramite Internet o una rete di comunicazione locale. Per ulteriori informazioni, consulta [Cos'è l'IoT?](#)

interpretabilità

Una caratteristica di un modello di machine learning che descrive il grado in cui un essere umano è in grado di comprendere in che modo le previsioni del modello dipendono dai suoi input. Per ulteriori informazioni, vedere Interpretabilità del modello di [machine learning](#) con AWS

IoT

Vedi [Internet of Things](#).

libreria di informazioni IT (ITIL)

Una serie di best practice per offrire servizi IT e allinearli ai requisiti aziendali. ITIL fornisce le basi per ITSM.

gestione dei servizi IT (ITSM)

Attività associate alla progettazione, implementazione, gestione e supporto dei servizi IT per un'organizzazione. Per informazioni sull'integrazione delle operazioni cloud con gli strumenti ITSM, consulta la [guida all'integrazione delle operazioni](#).

ITIL

Vedi la [libreria di informazioni IT](#).

ITSM

Vedi [Gestione dei servizi IT](#).

L

controllo degli accessi basato su etichette (LBAC)

Un'implementazione del controllo di accesso obbligatorio (MAC) in cui agli utenti e ai dati stessi viene assegnato esplicitamente un valore di etichetta di sicurezza. L'intersezione tra l'etichetta di sicurezza utente e l'etichetta di sicurezza dei dati determina quali righe e colonne possono essere visualizzate dall'utente.

zona di destinazione

Una landing zone è un AWS ambiente multi-account ben progettato, scalabile e sicuro. Questo è un punto di partenza dal quale le organizzazioni possono avviare e distribuire rapidamente carichi di lavoro e applicazioni con fiducia nel loro ambiente di sicurezza e infrastruttura. Per ulteriori informazioni sulle zone di destinazione, consulta la sezione [Configurazione di un ambiente AWS multi-account sicuro e scalabile](#).

modello linguistico di grandi dimensioni (LLM)

Un modello di [intelligenza artificiale](#) di deep learning preaddestrato su una grande quantità di dati. Un LLM può svolgere più attività, come rispondere a domande, riepilogare documenti, tradurre testo in altre lingue e completare frasi. [Per ulteriori informazioni, consulta Cosa sono. LLMs](#)

migrazione su larga scala

Una migrazione di 300 o più server.

BIANCO

Vedi controllo degli accessi [basato su etichette](#).

Privilegio minimo

La best practice di sicurezza per la concessione delle autorizzazioni minime richieste per eseguire un'attività. Per ulteriori informazioni, consulta [Applicazione delle autorizzazioni del privilegio minimo](#) nella documentazione di IAM.

eseguire il rehosting (lift and shift)

Vedi [7 R](#).

sistema little-endian

Un sistema che memorizza per primo il byte meno importante. Vedi anche [endianità](#).

LLM

Vedi modello [linguistico di grandi dimensioni](#).

ambienti inferiori

Vedi [ambiente](#).

M

machine learning (ML)

Un tipo di intelligenza artificiale che utilizza algoritmi e tecniche per il riconoscimento e l'apprendimento di schemi. Il machine learning analizza e apprende dai dati registrati, come i dati dell'Internet delle cose (IoT), per generare un modello statistico basato su modelli. Per ulteriori informazioni, consulta la sezione [Machine learning](#).

ramo principale

Vedi [filiale](#).

malware

Software progettato per compromettere la sicurezza o la privacy del computer. Il malware potrebbe interrompere i sistemi informatici, divulgare informazioni sensibili o ottenere accessi non autorizzati. Esempi di malware includono virus, worm, ransomware, trojan horse, spyware e keylogger.

servizi gestiti

Servizi AWS per cui AWS gestisce il livello di infrastruttura, il sistema operativo e le piattaforme e si accede agli endpoint per archiviare e recuperare i dati. Amazon Simple Storage Service (Amazon S3) Simple Storage Service (Amazon S3) e Amazon DynamoDB sono esempi di servizi gestiti. Questi sono noti anche come servizi astratti.

sistema di esecuzione della produzione (MES)

Un sistema software per tracciare, monitorare, documentare e controllare i processi di produzione che convertono le materie prime in prodotti finiti in officina.

MAP

Vedi [Migration Acceleration Program](#).

meccanismo

Un processo completo in cui si crea uno strumento, si promuove l'adozione dello strumento e quindi si esaminano i risultati per apportare le modifiche. Un meccanismo è un ciclo che si rafforza e si migliora man mano che funziona. Per ulteriori informazioni, consulta [Creazione di meccanismi nel AWS Well-Architected Framework](#).

account membro

Tutti gli account Account AWS diversi dall'account di gestione che fanno parte di un'organizzazione in. AWS Organizations Un account può essere membro di una sola organizzazione alla volta.

MEH

Vedi [sistema di esecuzione della produzione](#).

Message Queuing Telemetry Transport (MQTT)

[Un protocollo di comunicazione machine-to-machine \(M2M\) leggero, basato sul modello di pubblicazione/sottoscrizione, per dispositivi IoT con risorse limitate.](#)

microservizio

Un servizio piccolo e indipendente che comunica tramite canali ben definiti ed è in genere di proprietà di piccoli team autonomi. APIs Ad esempio, un sistema assicurativo potrebbe includere microservizi che si riferiscono a funzionalità aziendali, come vendite o marketing, o sottodomini, come acquisti, reclami o analisi. I vantaggi dei microservizi includono agilità, dimensionamento flessibile, facilità di implementazione, codice riutilizzabile e resilienza. Per ulteriori informazioni, consulta [Integrazione dei microservizi utilizzando servizi serverless](#). AWS

architettura di microservizi

Un approccio alla creazione di un'applicazione con componenti indipendenti che eseguono ogni processo applicativo come microservizio. Questi microservizi comunicano attraverso un'interfaccia ben definita utilizzando sistemi leggeri. APIs Ogni microservizio in questa architettura può essere aggiornato, distribuito e dimensionato per soddisfare la richiesta di funzioni specifiche di un'applicazione. Per ulteriori informazioni, vedere [Implementazione dei microservizi](#) su. AWS

Programma di accelerazione della migrazione (MAP)

Un AWS programma che fornisce consulenza, supporto, formazione e servizi per aiutare le organizzazioni a costruire una solida base operativa per il passaggio al cloud e per contribuire a compensare il costo iniziale delle migrazioni. MAP include una metodologia di migrazione per

eseguire le migrazioni precedenti in modo metodico e un set di strumenti per automatizzare e accelerare gli scenari di migrazione comuni.

migrazione su larga scala

Il processo di trasferimento della maggior parte del portfolio di applicazioni sul cloud avviene a ondate, con più applicazioni trasferite a una velocità maggiore in ogni ondata. Questa fase utilizza le migliori pratiche e le lezioni apprese nelle fasi precedenti per implementare una fabbrica di migrazione di team, strumenti e processi per semplificare la migrazione dei carichi di lavoro attraverso l'automazione e la distribuzione agile. Questa è la terza fase della [strategia di migrazione AWS](#).

fabbrica di migrazione

Team interfunzionali che semplificano la migrazione dei carichi di lavoro attraverso approcci automatizzati e agili. I team di Migration Factory includono in genere operazioni, analisti e proprietari aziendali, ingegneri addetti alla migrazione, sviluppatori e DevOps professionisti che lavorano nell'ambito degli sprint. Tra il 20% e il 50% di un portfolio di applicazioni aziendali è costituito da schemi ripetuti che possono essere ottimizzati con un approccio di fabbrica. Per ulteriori informazioni, consulta la [discussione sulle fabbriche di migrazione](#) e la [Guida alla fabbrica di migrazione al cloud](#) in questo set di contenuti.

metadati di migrazione

Le informazioni sull'applicazione e sul server necessarie per completare la migrazione. Ogni modello di migrazione richiede un set diverso di metadati di migrazione. Esempi di metadati di migrazione includono la sottorete, il gruppo di sicurezza e l'account di destinazione. AWS

modello di migrazione

Un'attività di migrazione ripetibile che descrive in dettaglio la strategia di migrazione, la destinazione della migrazione e l'applicazione o il servizio di migrazione utilizzati. Esempio: riorganizza la migrazione su Amazon EC2 AWS con Application Migration Service.

Valutazione del portfolio di migrazione (MPA)

Uno strumento online che fornisce informazioni per la convalida del business case per la migrazione a. Cloud AWS MPA offre una valutazione dettagliata del portfolio (dimensionamento corretto dei server, prezzi, confronto del TCO, analisi dei costi di migrazione) e pianificazione della migrazione (analisi e raccolta dei dati delle applicazioni, raggruppamento delle applicazioni, prioritizzazione delle migrazioni e pianificazione delle ondate). [Lo strumento MPA](#) (richiede l'accesso) è disponibile gratuitamente per tutti i AWS consulenti e i consulenti dei partner APN.

valutazione della preparazione alla migrazione (MRA)

Il processo di acquisizione di informazioni sullo stato di preparazione al cloud di un'organizzazione, l'identificazione dei punti di forza e di debolezza e la creazione di un piano d'azione per colmare le lacune identificate, utilizzando il CAF. AWS Per ulteriori informazioni, consulta la [guida di preparazione alla migrazione](#). MRA è la prima fase della [strategia di migrazione AWS](#).

strategia di migrazione

L'approccio utilizzato per migrare un carico di lavoro verso. Cloud AWS Per ulteriori informazioni, consulta la voce [7 R](#) in questo glossario e consulta [Mobilita la tua organizzazione per](#) accelerare le migrazioni su larga scala.

ML

[Vedi machine learning.](#)

modernizzazione

Trasformazione di un'applicazione obsoleta (legacy o monolitica) e della relativa infrastruttura in un sistema agile, elastico e altamente disponibile nel cloud per ridurre i costi, aumentare l'efficienza e sfruttare le innovazioni. Per ulteriori informazioni, vedere [Strategia per la modernizzazione delle applicazioni in](#). Cloud AWS

valutazione della preparazione alla modernizzazione

Una valutazione che aiuta a determinare la preparazione alla modernizzazione delle applicazioni di un'organizzazione, identifica vantaggi, rischi e dipendenze e determina in che misura l'organizzazione può supportare lo stato futuro di tali applicazioni. Il risultato della valutazione è uno schema dell'architettura di destinazione, una tabella di marcia che descrive in dettaglio le fasi di sviluppo e le tappe fondamentali del processo di modernizzazione e un piano d'azione per colmare le lacune identificate. Per ulteriori informazioni, vedere [Valutazione della preparazione alla modernizzazione per](#) le applicazioni in. Cloud AWS

applicazioni monolitiche (monoliti)

Applicazioni eseguite come un unico servizio con processi strettamente collegati. Le applicazioni monolitiche presentano diversi inconvenienti. Se una funzionalità dell'applicazione registra un picco di domanda, l'intera architettura deve essere dimensionata. L'aggiunta o il miglioramento delle funzionalità di un'applicazione monolitica diventa inoltre più complessa man mano che la base di codice cresce. Per risolvere questi problemi, puoi utilizzare un'architettura di microservizi. Per ulteriori informazioni, consulta la sezione [Scomposizione dei monoliti in microservizi](#).

MAPPA

Vedi [Migration Portfolio Assessment](#).

MQTT

Vedi [Message Queuing Telemetry Transport](#).

classificazione multiclasse

Un processo che aiuta a generare previsioni per più classi (prevedendo uno o più di due risultati). Ad esempio, un modello di machine learning potrebbe chiedere "Questo prodotto è un libro, un'auto o un telefono?" oppure "Quale categoria di prodotti è più interessante per questo cliente?"

infrastruttura mutabile

Un modello che aggiorna e modifica l'infrastruttura esistente per i carichi di lavoro di produzione. Per migliorare la coerenza, l'affidabilità e la prevedibilità, il AWS Well-Architected Framework consiglia l'uso di un'infrastruttura [immutabile](#) come best practice.

O

OAC

Vedi [Origin Access Control](#).

QUERCIA

Vedi [Origin Access Identity](#).

OCM

Vedi [gestione delle modifiche organizzative](#).

migrazione offline

Un metodo di migrazione in cui il carico di lavoro di origine viene eliminato durante il processo di migrazione. Questo metodo prevede tempi di inattività prolungati e viene in genere utilizzato per carichi di lavoro piccoli e non critici.

OI

Vedi [l'integrazione delle operazioni](#).

OLA

Vedi accordo a [livello operativo](#).

migrazione online

Un metodo di migrazione in cui il carico di lavoro di origine viene copiato sul sistema di destinazione senza essere messo offline. Le applicazioni connesse al carico di lavoro possono continuare a funzionare durante la migrazione. Questo metodo comporta tempi di inattività pari a zero o comunque minimi e viene in genere utilizzato per carichi di lavoro di produzione critici.

OPC-UA

Vedi [Open Process Communications - Unified Architecture](#).

Comunicazioni a processo aperto - Architettura unificata (OPC-UA)

Un protocollo di comunicazione machine-to-machine (M2M) per l'automazione industriale. OPC-UA fornisce uno standard di interoperabilità con schemi di crittografia, autenticazione e autorizzazione dei dati.

accordo a livello operativo (OLA)

Un accordo che chiarisce quali sono gli impegni reciproci tra i gruppi IT funzionali, a supporto di un accordo sul livello di servizio (SLA).

revisione della prontezza operativa (ORR)

Un elenco di domande e best practice associate che aiutano a comprendere, valutare, prevenire o ridurre la portata degli incidenti e dei possibili guasti. Per ulteriori informazioni, vedere [Operational Readiness Reviews \(ORR\)](#) nel Well-Architected AWS Framework.

tecnologia operativa (OT)

Sistemi hardware e software che interagiscono con l'ambiente fisico per controllare le operazioni, le apparecchiature e le infrastrutture industriali. Nella produzione, l'integrazione di sistemi OT e di tecnologia dell'informazione (IT) è un obiettivo chiave per le trasformazioni [dell'Industria 4.0](#).

integrazione delle operazioni (OI)

Il processo di modernizzazione delle operazioni nel cloud, che prevede la pianificazione, l'automazione e l'integrazione della disponibilità. Per ulteriori informazioni, consulta la [guida all'integrazione delle operazioni](#).

trail organizzativo

Un percorso creato da noi AWS CloudTrail che registra tutti gli eventi di un'organizzazione per tutti Account AWS . AWS Organizations Questo percorso viene creato in ogni Account AWS che

fa parte dell'organizzazione e tiene traccia dell'attività in ogni account. Per ulteriori informazioni, consulta [Creazione di un percorso per un'organizzazione](#) nella CloudTrail documentazione.

gestione del cambiamento organizzativo (OCM)

Un framework per la gestione di trasformazioni aziendali importanti e che comportano l'interruzione delle attività dal punto di vista delle persone, della cultura e della leadership. OCM aiuta le organizzazioni a prepararsi e passare a nuovi sistemi e strategie accelerando l'adozione del cambiamento, affrontando i problemi di transizione e promuovendo cambiamenti culturali e organizzativi. Nella strategia di AWS migrazione, questo framework si chiama accelerazione delle persone, a causa della velocità di cambiamento richiesta nei progetti di adozione del cloud. Per ulteriori informazioni, consultare la [Guida OCM](#).

controllo dell'accesso all'origine (OAC)

In CloudFront, un'opzione avanzata per limitare l'accesso per proteggere i contenuti di Amazon Simple Storage Service (Amazon S3). OAC supporta tutti i bucket S3 in generale Regioni AWS, la crittografia lato server con AWS KMS (SSE-KMS) e le richieste dinamiche e dirette al bucket S3.
PUT DELETE

identità di accesso origine (OAI)

Nel CloudFront, un'opzione per limitare l'accesso per proteggere i tuoi contenuti Amazon S3. Quando usi OAI, CloudFront crea un principale con cui Amazon S3 può autenticarsi. I principali autenticati possono accedere ai contenuti in un bucket S3 solo tramite una distribuzione specifica. CloudFront Vedi anche [OAC](#), che fornisce un controllo degli accessi più granulare e avanzato.

ORR

[Vedi la revisione della prontezza operativa.](#)

NON

Vedi la [tecnologia operativa](#).

VPC in uscita (egress)

In un'architettura AWS multi-account, un VPC che gestisce le connessioni di rete avviate dall'interno di un'applicazione. La [AWS Security Reference Architecture](#) consiglia di configurare l'account di rete con funzionalità in entrata, in uscita e di ispezione VPCs per proteggere l'interfaccia bidirezionale tra l'applicazione e Internet in generale.

P

limite delle autorizzazioni

Una policy di gestione IAM collegata ai principali IAM per impostare le autorizzazioni massime che l'utente o il ruolo possono avere. Per ulteriori informazioni, consulta [Limiti delle autorizzazioni](#) nella documentazione di IAM.

informazioni di identificazione personale (PII)

Informazioni che, se visualizzate direttamente o abbinate ad altri dati correlati, possono essere utilizzate per dedurre ragionevolmente l'identità di un individuo. Esempi di informazioni personali includono nomi, indirizzi e informazioni di contatto.

Informazioni che consentono l'identificazione personale degli utenti

Visualizza le [informazioni di identificazione personale](#).

playbook

Una serie di passaggi predefiniti che raccolgono il lavoro associato alle migrazioni, come l'erogazione delle funzioni operative principali nel cloud. Un playbook può assumere la forma di script, runbook automatici o un riepilogo dei processi o dei passaggi necessari per gestire un ambiente modernizzato.

PLC

Vedi [controllore logico programmabile](#).

PLM

Vedi la gestione [del ciclo di vita del prodotto](#).

policy

[Un oggetto in grado di definire le autorizzazioni \(vedi politica basata sull'identità\), specificare le condizioni di accesso \(vedi politicabasata sulle risorse\) o definire le autorizzazioni massime per tutti gli account di un'organizzazione in \(vedi politica di controllo dei servizi\). AWS Organizations](#)

persistenza poliglotta

Scelta indipendente della tecnologia di archiviazione di dati di un microservizio in base ai modelli di accesso ai dati e ad altri requisiti. Se i microservizi utilizzano la stessa tecnologia di archiviazione di dati, possono incontrare problemi di implementazione o registrare prestazioni scadenti. I microservizi vengono implementati più facilmente e ottengono prestazioni e scalabilità migliori se utilizzano l'archivio dati più adatto alle loro esigenze.

valutazione del portfolio

Un processo di scoperta, analisi e definizione delle priorità del portfolio di applicazioni per pianificare la migrazione. Per ulteriori informazioni, consulta la pagina [Valutazione della preparazione alla migrazione](#).

predicate

Una condizione di interrogazione che restituisce o, in genere, si trova in una clausola `true`. `false` WHERE

predicato pushdown

Una tecnica di ottimizzazione delle query del database che filtra i dati della query prima del trasferimento. Ciò riduce la quantità di dati che devono essere recuperati ed elaborati dal database relazionale e migliora le prestazioni delle query.

controllo preventivo

Un controllo di sicurezza progettato per impedire il verificarsi di un evento. Questi controlli sono la prima linea di difesa per impedire accessi non autorizzati o modifiche indesiderate alla rete. Per ulteriori informazioni, consulta [Controlli preventivi](#) in Implementazione dei controlli di sicurezza in AWS.

principale

Un'entità in AWS grado di eseguire azioni e accedere alle risorse. Questa entità è in genere un utente root per un Account AWS ruolo IAM o un utente. Per ulteriori informazioni, consulta Principali in [Termini e concetti dei ruoli](#) nella documentazione di IAM.

privacy fin dalla progettazione

Un approccio di ingegneria dei sistemi che tiene conto della privacy durante l'intero processo di sviluppo.

zone ospitate private

Un contenitore che contiene informazioni su come desideri che Amazon Route 53 risponda alle query DNS per un dominio e i relativi sottodomini all'interno di uno o più VPCs. Per ulteriori informazioni, consulta [Utilizzo delle zone ospitate private](#) nella documentazione di Route 53.

controllo proattivo

Un [controllo di sicurezza](#) progettato per impedire l'implementazione di risorse non conformi. Questi controlli analizzano le risorse prima del loro provisioning. Se la risorsa non è conforme al

controllo, non viene fornita. Per ulteriori informazioni, consulta la [guida di riferimento sui controlli](#) nella AWS Control Tower documentazione e consulta Controlli [proattivi in Implementazione dei controlli](#) di sicurezza su. AWS

gestione del ciclo di vita del prodotto (PLM)

La gestione dei dati e dei processi di un prodotto durante l'intero ciclo di vita, dalla progettazione, sviluppo e lancio, attraverso la crescita e la maturità, fino al declino e alla rimozione.

Ambiente di produzione

[Vedi ambiente.](#)

controllore logico programmabile (PLC)

Nella produzione, un computer altamente affidabile e adattabile che monitora le macchine e automatizza i processi di produzione.

concatenamento rapido

Utilizzo dell'output di un prompt [LLM](#) come input per il prompt successivo per generare risposte migliori. Questa tecnica viene utilizzata per suddividere un'attività complessa in sottoattività o per perfezionare o espandere iterativamente una risposta preliminare. Aiuta a migliorare l'accuratezza e la pertinenza delle risposte di un modello e consente risultati più granulari e personalizzati.

pseudonimizzazione

Il processo di sostituzione degli identificatori personali in un set di dati con valori segnaposto. La pseudonimizzazione può aiutare a proteggere la privacy personale. I dati pseudonimizzati sono ancora considerati dati personali.

publish/subscribe (pub/sub)

Un modello che consente comunicazioni asincrone tra microservizi per migliorare la scalabilità e la reattività. Ad esempio, in un [MES](#) basato su microservizi, un microservizio può pubblicare messaggi di eventi su un canale a cui altri microservizi possono abbonarsi. Il sistema può aggiungere nuovi microservizi senza modificare il servizio di pubblicazione.

Q

Piano di query

Una serie di passaggi, come le istruzioni, utilizzati per accedere ai dati in un sistema di database relazionale SQL.

regressione del piano di query

Quando un ottimizzatore del servizio di database sceglie un piano non ottimale rispetto a prima di una determinata modifica all'ambiente di database. Questo può essere causato da modifiche a statistiche, vincoli, impostazioni dell'ambiente, associazioni dei parametri di query e aggiornamenti al motore di database.

R

Matrice RACI

Vedi [responsabile, responsabile, consultato, informato \(RACI\)](#).

RAG

Vedi [Retrieval](#) Augmented Generation.

ransomware

Un software dannoso progettato per bloccare l'accesso a un sistema informatico o ai dati fino a quando non viene effettuato un pagamento.

Matrice RASCI

Vedi [responsabile, responsabile, consultato, informato \(RACI\)](#).

RCAC

Vedi controllo dell'[accesso a righe e colonne](#).

replica di lettura

Una copia di un database utilizzata per scopi di sola lettura. È possibile indirizzare le query alla replica di lettura per ridurre il carico sul database principale.

riprogettare

Vedi [7 Rs](#).

obiettivo del punto di ripristino (RPO)

Il periodo di tempo massimo accettabile dall'ultimo punto di ripristino dei dati. Questo determina ciò che si considera una perdita di dati accettabile tra l'ultimo punto di ripristino e l'interruzione del servizio.

obiettivo del tempo di ripristino (RTO)

Il ritardo massimo accettabile tra l'interruzione del servizio e il ripristino del servizio.

rifattorizzare

Vedi [7 R.](#)

Region

Una raccolta di AWS risorse in un'area geografica. Ciascuna Regione AWS è isolata e indipendente dalle altre per fornire tolleranza agli errori, stabilità e resilienza. Per ulteriori informazioni, consulta [Specificare cosa può usare Regioni AWS il tuo account.](#)

regressione

Una tecnica di ML che prevede un valore numerico. Ad esempio, per risolvere il problema "A che prezzo verrà venduta questa casa?" un modello di ML potrebbe utilizzare un modello di regressione lineare per prevedere il prezzo di vendita di una casa sulla base di dati noti sulla casa (ad esempio, la metratura).

riospitare

Vedi [7 R.](#)

rilascio

In un processo di implementazione, l'atto di promuovere modifiche a un ambiente di produzione.

trasferisco

Vedi [7 Rs.](#)

ripiattaforma

Vedi [7 Rs.](#)

riacquisto

Vedi [7 Rs.](#)

resilienza

La capacità di un'applicazione di resistere alle interruzioni o di ripristinarle. [L'elevata disponibilità e il disaster recovery](#) sono considerazioni comuni quando si pianifica la resilienza in Cloud AWS. [Per ulteriori informazioni, vedere Cloud AWS Resilience.](#)

policy basata su risorse

Una policy associata a una risorsa, ad esempio un bucket Amazon S3, un endpoint o una chiave di crittografia. Questo tipo di policy specifica a quali principali è consentito l'accesso, le azioni supportate e qualsiasi altra condizione che deve essere soddisfatta.

matrice di assegnazione di responsabilità (RACI)

Una matrice che definisce i ruoli e le responsabilità di tutte le parti coinvolte nelle attività di migrazione e nelle operazioni cloud. Il nome della matrice deriva dai tipi di responsabilità definiti nella matrice: responsabile (R), responsabile (A), consultato (C) e informato (I). Il tipo di supporto (S) è facoltativo. Se includi il supporto, la matrice viene chiamata matrice RASCI e, se la escludi, viene chiamata matrice RACI.

controllo reattivo

Un controllo di sicurezza progettato per favorire la correzione di eventi avversi o deviazioni dalla baseline di sicurezza. Per ulteriori informazioni, consulta [Controlli reattivi](#) in Implementazione dei controlli di sicurezza in AWS.

retain

Vedi [7 R](#).

andare in pensione

Vedi [7 Rs](#).

Retrieval Augmented Generation (RAG)

Una tecnologia di [intelligenza artificiale generativa](#) in cui un [LLM](#) fa riferimento a una fonte di dati autorevole esterna alle sue fonti di dati di formazione prima di generare una risposta. Ad esempio, un modello RAG potrebbe eseguire una ricerca semantica nella knowledge base o nei dati personalizzati di un'organizzazione. Per ulteriori informazioni, consulta [Cos'è il RAG](#).

rotazione

Processo di aggiornamento periodico di un [segreto](#) per rendere più difficile l'accesso alle credenziali da parte di un utente malintenzionato.

controllo dell'accesso a righe e colonne (RCAC)

L'uso di espressioni SQL di base e flessibili con regole di accesso definite. RCAC è costituito da autorizzazioni di riga e maschere di colonna.

RPO

Vedi [obiettivo del punto di ripristino](#).

VERSO

Vedi [obiettivo del tempo di ripristino](#).

runbook

Un insieme di procedure manuali o automatizzate necessarie per eseguire un'attività specifica. In genere sono progettati per semplificare operazioni o procedure ripetitive con tassi di errore elevati.

S

SAML 2.0

Uno standard aperto utilizzato da molti provider di identità (IdPs). Questa funzionalità abilita il single sign-on (SSO) federato, in modo che gli utenti possano accedere Console di gestione AWS o chiamare le operazioni AWS API senza che tu debba creare un utente in IAM per tutti i membri dell'organizzazione. Per ulteriori informazioni sulla federazione basata su SAML 2.0, consulta [Informazioni sulla federazione basata su SAML 2.0](#) nella documentazione di IAM.

SCADA

Vedi [controllo di supervisione e acquisizione dati](#).

SCP

Vedi la [politica di controllo del servizio](#).

Secret

In Gestione dei segreti AWS, informazioni riservate o riservate, come una password o le credenziali utente, archiviate in forma crittografata. È costituito dal valore segreto e dai relativi metadati. Il valore segreto può essere binario, una stringa singola o più stringhe. Per ulteriori informazioni, consulta [Cosa c'è in un segreto di Secrets Manager?](#) nella documentazione di Secrets Manager.

sicurezza fin dalla progettazione

Un approccio di ingegneria dei sistemi che tiene conto della sicurezza durante l'intero processo di sviluppo.

controllo di sicurezza

Un guardrail tecnico o amministrativo che impedisce, rileva o riduce la capacità di un autore di minacce di sfruttare una vulnerabilità di sicurezza. [Esistono quattro tipi principali di controlli di sicurezza: preventivi, investigativi, reattivi e proattivi.](#)

rafforzamento della sicurezza

Il processo di riduzione della superficie di attacco per renderla più resistente agli attacchi. Può includere azioni come la rimozione di risorse che non sono più necessarie, l'implementazione di best practice di sicurezza che prevedono la concessione del privilegio minimo o la disattivazione di funzionalità non necessarie nei file di configurazione.

sistema di gestione delle informazioni e degli eventi di sicurezza (SIEM)

Strumenti e servizi che combinano sistemi di gestione delle informazioni di sicurezza (SIM) e sistemi di gestione degli eventi di sicurezza (SEM). Un sistema SIEM raccoglie, monitora e analizza i dati da server, reti, dispositivi e altre fonti per rilevare minacce e violazioni della sicurezza e generare avvisi.

automazione della risposta alla sicurezza

Un'azione predefinita e programmata progettata per rispondere o porre rimedio automaticamente a un evento di sicurezza. Queste automazioni fungono da controlli di sicurezza [investigativi](#) o [reattivi](#) che aiutano a implementare le migliori pratiche di sicurezza. AWS Esempi di azioni di risposta automatizzate includono la modifica di un gruppo di sicurezza VPC, l'applicazione di patch a un'istanza Amazon EC2 o la rotazione delle credenziali.

Crittografia lato server

Crittografia dei dati a destinazione, da parte di chi li riceve. Servizio AWS

Policy di controllo dei servizi (SCP)

Una politica che fornisce il controllo centralizzato sulle autorizzazioni per tutti gli account di un'organizzazione in. AWS Organizations SCPs definire barriere o fissare limiti alle azioni che un amministratore può delegare a utenti o ruoli. È possibile utilizzarli SCPs come elenchi consentiti o elenchi di rifiuto, per specificare quali servizi o azioni sono consentiti o proibiti. Per ulteriori informazioni, consulta [le politiche di controllo del servizio](#) nella AWS Organizations documentazione.

endpoint del servizio

L'URL del punto di ingresso per un Servizio AWS. Puoi utilizzare l'endpoint per connetterti a livello di programmazione al servizio di destinazione. Per ulteriori informazioni, consulta [Endpoint del Servizio AWS](#) nei Riferimenti generali di AWS.

accordo sul livello di servizio (SLA)

Un accordo che chiarisce ciò che un team IT promette di offrire ai propri clienti, ad esempio l'operatività e le prestazioni del servizio.

indicatore del livello di servizio (SLI)

Misurazione di un aspetto prestazionale di un servizio, ad esempio il tasso di errore, la disponibilità o la velocità effettiva.

obiettivo a livello di servizio (SLO)

[Una metrica target che rappresenta lo stato di un servizio, misurato da un indicatore del livello di servizio.](#)

Modello di responsabilità condivisa

Un modello che descrive la responsabilità condivisa AWS per la sicurezza e la conformità del cloud. AWS è responsabile della sicurezza del cloud, mentre tu sei responsabile della sicurezza nel cloud. Per ulteriori informazioni, consulta [Modello di responsabilità condivisa](#).

SIEM

Vedi il [sistema di gestione delle informazioni e degli eventi sulla sicurezza](#).

punto di errore singolo (SPOF)

Un guasto in un singolo componente critico di un'applicazione che può disturbare il sistema.

SLAM

Vedi il contratto sul [livello di servizio](#).

SLI

Vedi l'indicatore del [livello di servizio](#).

LENTA

Vedi obiettivo del [livello di servizio](#).

split-and-seed modello

Un modello per dimensionare e accelerare i progetti di modernizzazione. Man mano che vengono definite nuove funzionalità e versioni dei prodotti, il team principale si divide per creare nuovi team di prodotto. Questo aiuta a dimensionare le capacità e i servizi dell'organizzazione, migliora la produttività degli sviluppatori e supporta una rapida innovazione. Per ulteriori informazioni, vedere [Approccio graduale alla modernizzazione delle applicazioni in](#). Cloud AWS

SPOF

Vedi [punto di errore singolo](#).

schema a stella

Una struttura organizzativa di database che utilizza un'unica tabella dei fatti di grandi dimensioni per archiviare i dati transazionali o misurati e utilizza una o più tabelle dimensionali più piccole per memorizzare gli attributi dei dati. Questa struttura è progettata per l'uso in un [data warehouse](#) o per scopi di business intelligence.

modello del fico strangolatore

Un approccio alla modernizzazione dei sistemi monolitici mediante la riscrittura e la sostituzione incrementali delle funzionalità del sistema fino alla disattivazione del sistema legacy. Questo modello utilizza l'analogia di una pianta di fico che cresce fino a diventare un albero robusto e alla fine annienta e sostituisce il suo ospite. Il modello è stato [introdotto da Martin Fowler](#) come metodo per gestire il rischio durante la riscrittura di sistemi monolitici. Per un esempio di come applicare questo modello, consulta [Modernizzazione incrementale dei servizi Web legacy di Microsoft ASP.NET \(ASMX\) mediante container e Gateway Amazon API](#).

sottorete

Un intervallo di indirizzi IP nel VPC. Una sottorete deve risiedere in una singola zona di disponibilità.

controllo di supervisione e acquisizione dati (SCADA)

Nella produzione, un sistema che utilizza hardware e software per monitorare gli asset fisici e le operazioni di produzione.

crittografia simmetrica

Un algoritmo di crittografia che utilizza la stessa chiave per crittografare e decrittografare i dati.

test sintetici

Test di un sistema in modo da simulare le interazioni degli utenti per rilevare potenziali problemi o monitorare le prestazioni. Puoi usare [Amazon CloudWatch Synthetics](#) per creare questi test.

prompt di sistema

Una tecnica per fornire contesto, istruzioni o linee guida a un [LLM](#) per indirizzarne il comportamento. I prompt di sistema aiutano a impostare il contesto e stabilire regole per le interazioni con gli utenti.

T

tag

Coppie chiave-valore che fungono da metadati per l'organizzazione delle risorse. AWS Con i tag è possibile a gestire, identificare, organizzare, cercare e filtrare le risorse. Per ulteriori informazioni, consulta [Tagging delle risorse AWS](#).

variabile di destinazione

Il valore che stai cercando di prevedere nel machine learning supervisionato. Questo è indicato anche come variabile di risultato. Ad esempio, in un ambiente di produzione la variabile di destinazione potrebbe essere un difetto del prodotto.

elenco di attività

Uno strumento che viene utilizzato per tenere traccia dei progressi tramite un runbook. Un elenco di attività contiene una panoramica del runbook e un elenco di attività generali da completare. Per ogni attività generale, include la quantità stimata di tempo richiesta, il proprietario e lo stato di avanzamento.

ambiente di test

[Vedi ambiente.](#)

training

Fornire dati da cui trarre ispirazione dal modello di machine learning. I dati di training devono contenere la risposta corretta. L'algoritmo di apprendimento trova nei dati di addestramento i pattern che mappano gli attributi dei dati di input al target (la risposta che si desidera prevedere). Produce un modello di ML che acquisisce questi modelli. Puoi quindi utilizzare il modello di ML per creare previsioni su nuovi dati di cui non si conosce il target.

Transit Gateway

Un hub di transito di rete che puoi utilizzare per interconnettere le tue reti VPCs e quelle locali. Per ulteriori informazioni, consulta [Cos'è un gateway di transito](#) nella AWS Transit Gateway documentazione.

flusso di lavoro basato su trunk

Un approccio in cui gli sviluppatori creano e testano le funzionalità localmente in un ramo di funzionalità e quindi uniscono tali modifiche al ramo principale. Il ramo principale viene quindi integrato negli ambienti di sviluppo, preproduzione e produzione, in sequenza.

Accesso attendibile

Concessione delle autorizzazioni a un servizio specificato dall'utente per eseguire attività all'interno dell'organizzazione AWS Organizations e nei suoi account per conto dell'utente. Il servizio attendibile crea un ruolo collegato al servizio in ogni account, quando tale ruolo è necessario, per eseguire attività di gestione per conto dell'utente. Per ulteriori informazioni, consulta [Utilizzo AWS Organizations con altri AWS servizi](#) nella AWS Organizations documentazione.

regolazione

Modificare alcuni aspetti del processo di training per migliorare la precisione del modello di ML. Ad esempio, puoi addestrare il modello di ML generando un set di etichette, aggiungendo etichette e quindi ripetendo questi passaggi più volte con impostazioni diverse per ottimizzare il modello.

team da due pizze

Una piccola DevOps squadra che puoi sfamare con due pizze. Un team composto da due persone garantisce la migliore opportunità possibile di collaborazione nello sviluppo del software.

U

incertezza

Un concetto che si riferisce a informazioni imprecise, incomplete o sconosciute che possono minare l'affidabilità dei modelli di machine learning predittivi. Esistono due tipi di incertezza: l'incertezza epistemica, che è causata da dati limitati e incompleti, mentre l'incertezza aleatoria è causata dal rumore e dalla casualità insiti nei dati.

compiti indifferenziati

Conosciuto anche come sollevamento di carichi pesanti, è un lavoro necessario per creare e far funzionare un'applicazione, ma che non apporta valore diretto all'utente finale né offre vantaggi competitivi. Esempi di attività indifferenziate includono l'approvvigionamento, la manutenzione e la pianificazione della capacità.

ambienti superiori

[Vedi ambiente.](#)

V

vacuum

Un'operazione di manutenzione del database che prevede la pulizia dopo aggiornamenti incrementali per recuperare lo spazio di archiviazione e migliorare le prestazioni.

controllo delle versioni

Processi e strumenti che tengono traccia delle modifiche, ad esempio le modifiche al codice di origine in un repository.

Peering VPC

Una connessione tra due VPCs che consente di indirizzare il traffico utilizzando indirizzi IP privati. Per ulteriori informazioni, consulta [Che cos'è il peering VPC?](#) nella documentazione di Amazon VPC.

vulnerabilità

Un difetto software o hardware che compromette la sicurezza del sistema.

W

cache calda

Una cache del buffer che contiene dati correnti e pertinenti a cui si accede frequentemente. L'istanza di database può leggere dalla cache del buffer, il che richiede meno tempo rispetto alla lettura dalla memoria dal disco principale.

dati caldi

Dati a cui si accede raramente. Quando si eseguono interrogazioni di questo tipo di dati, in genere sono accettabili query moderatamente lente.

funzione finestra

Una funzione SQL che esegue un calcolo su un gruppo di righe che si riferiscono in qualche modo al record corrente. Le funzioni della finestra sono utili per l'elaborazione di attività, come il calcolo di una media mobile o l'accesso al valore delle righe in base alla posizione relativa della riga corrente.

Carico di lavoro

Una raccolta di risorse e codice che fornisce valore aziendale, ad esempio un'applicazione rivolta ai clienti o un processo back-end.

flusso di lavoro

Gruppi funzionali in un progetto di migrazione responsabili di una serie specifica di attività. Ogni flusso di lavoro è indipendente ma supporta gli altri flussi di lavoro del progetto. Ad esempio, il flusso di lavoro del portfolio è responsabile della definizione delle priorità delle applicazioni, della pianificazione delle ondate e della raccolta dei metadati di migrazione. Il flusso di lavoro del portfolio fornisce queste risorse al flusso di lavoro di migrazione, che quindi migra i server e le applicazioni.

VERME

Vedi [scrivere una volta, leggere molti](#).

WQF

Vedi [AWS Workload Qualification Framework](#).

scrivi una volta, leggi molte (WORM)

Un modello di storage che scrive i dati una sola volta e ne impedisce l'eliminazione o la modifica. Gli utenti autorizzati possono leggere i dati tutte le volte che è necessario, ma non possono modificarli. Questa infrastruttura di archiviazione dei dati è considerata [immutabile](#).

Z

exploit zero-day

[Un attacco, in genere malware, che sfrutta una vulnerabilità zero-day.](#)

vulnerabilità zero-day

Un difetto o una vulnerabilità assoluta in un sistema di produzione. Gli autori delle minacce possono utilizzare questo tipo di vulnerabilità per attaccare il sistema. Gli sviluppatori vengono spesso a conoscenza della vulnerabilità causata dall'attacco.

prompt zero-shot

Fornire a un [LLM](#) le istruzioni per eseguire un'attività ma non esempi (immagini) che possano aiutarla. Il LLM deve utilizzare le sue conoscenze pre-addestrate per gestire l'attività. L'efficacia del prompt zero-shot dipende dalla complessità dell'attività e dalla qualità del prompt. [Vedi anche few-shot prompting.](#)

applicazione zombie

Un'applicazione che prevede un utilizzo CPU e memoria inferiore al 5%. In un progetto di migrazione, è normale ritirare queste applicazioni.

Le traduzioni sono generate tramite traduzione automatica. In caso di conflitto tra il contenuto di una traduzione e la versione originale in Inglese, quest'ultima prevarrà.