

L'architettura Data Fabric offre 3 vantaggi immediati

Integra e utilizza tutti i dati della
tua organizzazione per migliorare i
risultati finali di business

Indice

- 3** Le sfide che i dati propongono oggi
- 5** Vantaggi di un'architettura Data Fabric per il business
- 6** In che modo IBM fornisce un'architettura Data Fabric?
- 7** Data Fabric in azione
- 7** Data Fabric è il prossimo passo per la trasformazione della la tua azienda

Le sfide che i dati propongono oggi

I dati sono un elemento integrante della trasformazione digitale per le aziende.

Ma, quando le organizzazioni cercano di ricavare il massimo dai loro dati, affrontano sfide derivanti dalla varietà della provenienza dei dati stessi, tipi, strutture, ambienti e piattaforme. Questa difficile situazione di dati multidimensionali si complica ulteriormente quando le organizzazioni adottano architetture ibride e multicloud. Per molte aziende oggi, i dati operativi sono rimasti in gran parte isolati in silos e nascosti, con il risultato di un'enorme quantità di dati oscuri.

Prendiamo l'esempio di un'azienda energetica nordamericana che ha cercato di reinventarsi come organizzazione basata sui dati, per diventare un'azienda in cui le funzionalità della data science fossero facilmente accessibili in più unità di business. Si sono presto resi conto che la loro trasformazione digitale era ostacolata da dati isolati in silos, strumenti incoerenti e livelli di competenze diversi, tutti fattori che causavano divari critici tra le competenze in materia di dati.

Il problema che hanno affrontato non era specifico del loro business; infatti è una conseguenza comune dei panorami di dati la cui crescita ha superato le rispettive architetture di gestione.

Ciò che alla fine ha riportato l'azienda energetica sulla strada giusta per ottenere il successo nella trasformazione digitale è stato l'impiego di un nuovo concetto di architettura dei dati noto come Data Fabric.

Cos'è esattamente un'architettura Data Fabric, in cosa si differenzia dalle precedenti architetture, quali vantaggi può apportare alle aziende e quale ruolo svolge IBM nell'implementazione di tale architettura? In questo white paper si troveranno le risposte a tali domande.

Cosa si intende per architettura Data Fabric?

In passato, le organizzazioni hanno tentato di affrontare i problemi di accesso ai dati attraverso l'integrazione punto a punto o l'introduzione di hub di dati. Nessuno di questi due approcci è adatto quando i dati sono estremamente distribuiti e isolati in silos. Le integrazioni punto a punto aggiungono un costo esponenziale per ogni endpoint aggiuntivo che dovrà essere collegato, il che significa che si tratta di un approccio che non fornisce scalabilità. Gli hub di dati consentono un'integrazione più facile di applicazioni e origini dati, ma aumentano il costo e la complessità quando si tratta di riuscire a garantire qualità e attendibilità dei dati all'interno dell'hub.

Data Fabric è un'architettura emergente che mira ad affrontare le sfide imposte dai dati e che derivano da un panorama di dati ibridi. L'idea su cui si fonda è quella di trovare un equilibrio tra decentralizzazione e globalizzazione, fungendo da tessuto connettivo virtuale tra gli endpoint dei dati (vedere pagina 4).

Attraverso tecnologie come l'automazione e l'aumento dell'integrazione, la governance federata e l'attivazione dei metadati, un'architettura Data Fabric permette l'orchestrazione dinamica e intelligente dei dati in un panorama distribuito, creando una rete di informazioni immediatamente disponibili da utilizzare al meglio in un'azienda.

Un'architettura Data Fabric è indipendente dalle piattaforme di distribuzione, dai processi dei dati, dalle ubicazioni geografiche e dall'approccio architettonico. Facilita l'uso dei dati come risorsa aziendale. Un'architettura Data Fabric assicura che i vari tipi di dati possano essere combinati, visualizzati e regolamentati correttamente, in modo efficiente ed efficace.

Funzionalità e principi di un'architettura Data Fabric

Alla base dell'architettura Data Fabric c'è una piattaforma di gestione dei dati che mette a disposizione l'intera gamma di funzionalità di gestione integrata dei dati, che includono rilevamento, governance, cura e orchestrazione.

Tuttavia, un'architettura Data Fabric rappresenta un avanzamento e un'evoluzione rispetto ai concetti tradizionali di gestione dei dati, quali ad esempio DataOps, che si focalizza solo sullo stabilire procedure per aumentare il livello di operazionalizzazione dei dati. Questo approccio si basa su un'architettura distribuita e una tecnologia avanzata, in grado di affrontare le esigenze che nascono dall'estrema diversità e distribuzione delle risorse di dati.

Un'architettura Data Fabric potrebbe essere logicamente suddivisa in quattro funzionalità (o componenti):

Conoscenze, insight e semantica

- Si può paragonare a un mercato di dati e un'esperienza di acquisto
- Arricchisce automaticamente le risorse di dati rilevate aggiungendo conoscenze e semantica, permettendo ai consumatori di trovare e comprendere i dati

Governance e conformità unificate

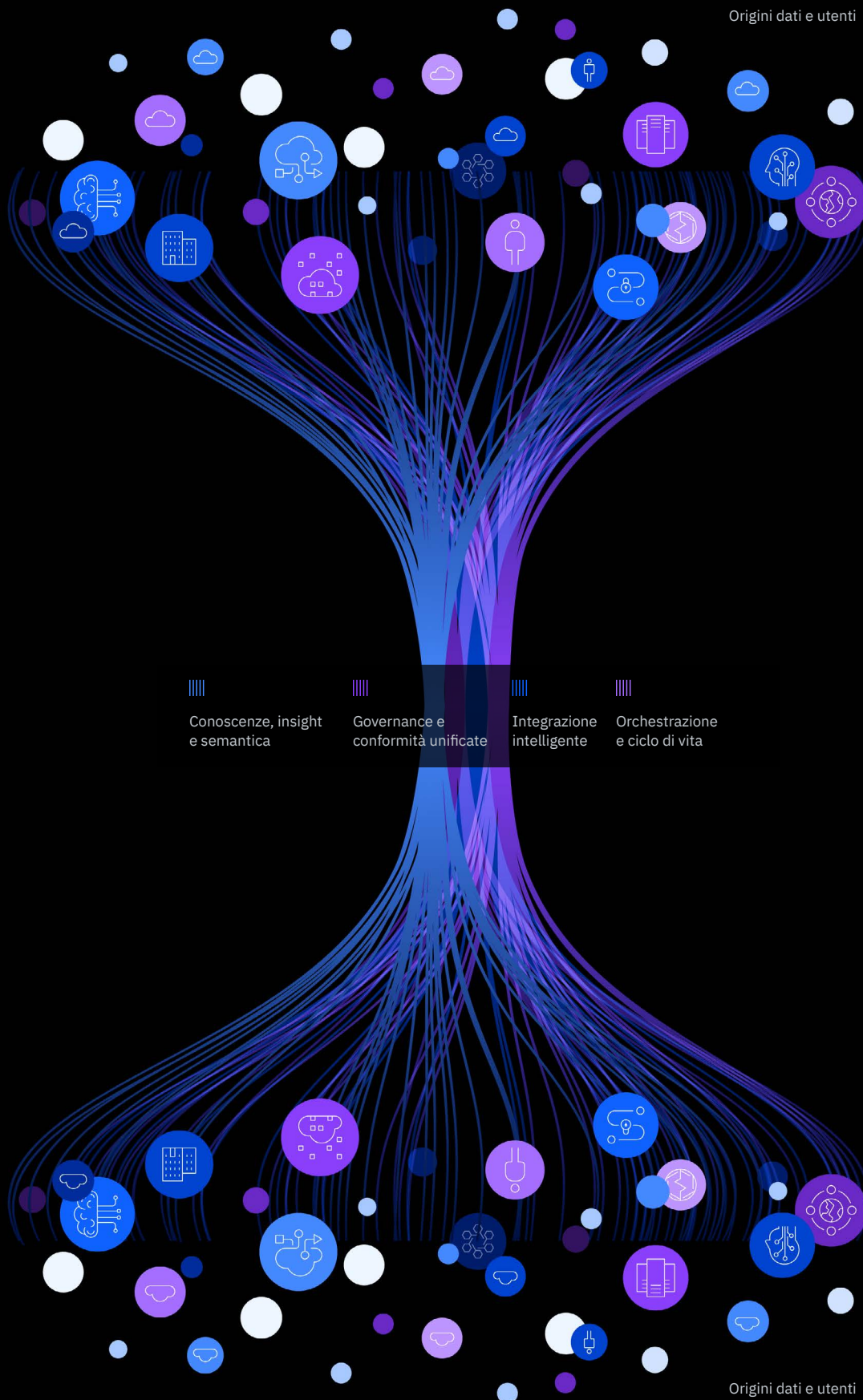
- Consente la gestione e la governance dei metadati a livello locale, ma supporta l'applicazione di politiche e una vista unificata a livello globale
- Applica automaticamente politiche alle risorse di dati, in conformità alle regole globali e locali
- Utilizza funzionalità avanzate per automatizzare la classificazione e la cura delle risorse di dati
- Stabilisce automaticamente percorsi di accesso disponibili per query per qualsiasi risorsa catalogata per una maggiore attivazione dei dati

Integrazione intelligente

- Accelera le attività di un ingegnere dei dati attraverso la creazione di flussi e pipeline automatici in origini dati distribuite
- Consente l'inserimento self-service e l'accesso ai dati per qualsiasi dato, con l'applicazione completa, a livello locale e globale, di politiche di protezione dei dati
- Determina automaticamente l'esecuzione più adatta attraverso la distribuzione ottimizzata del carico di lavoro, la regolazione automatica e la correzione delle deviazioni dallo schema

Orchestratura e ciclo di vita

- Abilita la composizione, il test, l'operatività e il monitoraggio delle pipeline di dati
- Incorpora funzionalità di AI nel ciclo di vita dei dati per automatizzare le attività, regolare e correggere automaticamente, nonché rilevare le modifiche dei dati di origine e tutto questo facilita gli aggiornamenti automatici



Vantaggi di un'architettura Data Fabric per il business

I dati generano valore di business solo quando sono contestualizzati e diventano accessibili da qualsiasi utente o applicazione nell'organizzazione. Se implementata correttamente, un'architettura Data Fabric aiuta a garantire che quei valori siano disponibili in tutta l'organizzazione nel modo più efficiente e automatizzato possibile. Come tale, l'architettura Data Fabric presenta tre vantaggi fondamentali:

1. Consentire un utilizzo dei dati e una collaborazione self-service
2. Automatizzare la governance, la protezione e la sicurezza; vantaggio realizzato mediante i metadati attivi
3. Automatizzare le attività di ingegneria dei dati e aumentare l'integrazione dei dati nelle risorse cloud ibride

Consentire un utilizzo dei dati e una collaborazione self-service

Integrando i dati provenienti da più origini e analizzando una frazione più estesa dell'enorme quantità di dati generata quotidianamente, le organizzazioni ottengono insight più validi e rispondono più rapidamente alle richieste variabili del business. Un'architettura Data Fabric distribuisce rapidamente i dati a coloro che ne hanno bisogno. La modalità self-service permette all'organizzazione nel suo insieme di individuare più rapidamente i dati appropriati e di dedicare più tempo all'utilizzo di quei dati per fornire insight concreti.

Vantaggi dell'architettura Data Fabric per l'utilizzo self-service dei dati:

- Gli utenti di business potranno individuare, comprendere, utilizzare i dati e dargli forma in tutta l'organizzazione.
- Una governance e una derivazione centralizzate dei dati aiutano gli utenti a comprendere il significato dei dati, la loro provenienza e il loro collegamento ad altre risorse.
- La gestione completa e personalizzabile dei metadati offre una facile scalabilità e accessibilità tramite API.
- L'accesso self-service a dati affidabili e regolamentati permette la collaborazione della LoB (line-of-business) con altri utenti.

Uno studio Forrester Total Economic Impact¹ ha rivelato che queste funzionalità possono significare:

USD
5,8
milioni 459%
di vantaggi ROI

Automatizzare la governance, la protezione dei dati e la sicurezza; vantaggio realizzato mediante i metadati attivi

Un livello di governance attiva e distribuita per tutte le iniziative relative ai dati riduce i rischi legati alla conformità e alla regolamentazione garantendo affidabilità e trasparenza.

Consente l'applicazione automatica delle politiche per l'accesso a qualsiasi dato, fornendo un elevato livello di protezione e conformità dei dati.

L'utilizzo delle tecnologie di AI e apprendimento automatico permette agli utenti dell'architettura Data Fabric di aumentare il proprio livello di automazione, ad esempio estraendo automaticamente le regole di governance dei dati sulla base della lingua e delle definizioni nei documenti normativi. Questo permette alle organizzazioni di applicare regole di governance specifiche del settore d'industria in pochi minuti per poter evitare multe costose e garantire un uso etico dei dati, ovunque essi risiedano.

Vantaggi di un'architettura Data Fabric per la virtualizzazione regolamentata:

- Per gli ingegneri dei dati, i data scientist e gli analisti di business sono aumentate agilità, sicurezza e produttività.
- Più origini dati globali appaiono come un unico database.
- Una nuova modalità di rilevamento, leader di settore, di PII (personally identifiable information - informazioni per l'identificazione personale) ed elementi critici dei dati è possibile su larghissima scala.

Queste funzionalità possono realizzare quanto segue:

USD
2,4
milioni 430%
di vantaggi¹ miglioramento delle prestazioni²

Automatizzare le attività di ingegneria dei dati e aumentare l'integrazione dei dati

L'ingegneria avanzata dei dati fa sì che praticamente qualsiasi processo di visualizzazione o distribuzione dei dati sia automatizzato e non richieda alcun processo di codifica noioso o soggetto a errori. L'aumento dell'integrazione utilizza i dati dei metadati per ottimizzare la distribuzione dei dati e l'accesso ad essi.

Vantaggi di un'architettura Data Fabric per l'ingegneria e l'integrazione dei dati:

- L'integrazione dei dati ottimizzata automaticamente consente di accelerare la distribuzione dei dati.
- Il bilanciamento automatico del carico di lavoro e la scalabilità elastica rendono i lavori pronti per qualsiasi ambiente e qualsiasi volume di dati.
- La resilienza e l'automazione del flusso CI/CD (integrazione e distribuzione continue) sono funzionalità integrate.
- Il processo automatizzato di acquisizione delle modifiche in tempo reale supporta la distribuzione di dati di qualità per i processi di business.
- L'apprendimento automatico è in grado di automatizzare ed estendere i processi di rilevamento, classificazione e cura dei dati personalizzati, portando a un time-to-value più veloce.
- L'analisi continua può essere eseguita automaticamente, in tempo reale, ovunque risiedano i dati.

I risultati di un rivenditore leader³ dimostrano:

60 volte 20 volte
accelerazione nei tempi di maggiore velocità nell'analisi
distribuzione dei dati dell'affinità dei clienti

In che modo IBM fornisce un'architettura Data Fabric?

IBM Cloud Pak® for Data rende possibile questo concetto di Data Fabric. IBM Cloud Pak for Data è una piattaforma di insight che semplifica e automatizza la raccolta, l'organizzazione e l'analisi dei dati e accelera l'incorporazione dell'AI in tutta l'azienda.

Con le sue funzionalità che consentono di collegare i dati ovunque; eseguire i carichi di lavoro ovunque e creare, distribuire e gestire l'AI su larga scala in ambienti di cloud ibrido, IBM cloud Pak for Data è l'abilitatore per una trasformazione digitale del business.

La piattaforma offre un'integrazione senza soluzione di continuità in un ambiente aziendale ibrido per:

- Servizi IBM Cloud Pak for Data
- Applicazioni e origini dati esterne
- Funzionalità avanzate basate sull'AI per la gestione e la governance dei dati

Tramite questa base si mettono a disposizione dei consumatori dati curati, con un equilibrio ottimale di costi, prestazioni, conformità e con l'intelligence per orchestrare e ottimizzare l'elaborazione dei dati in base ai carichi di lavoro, all'ubicazione dei dati e alle politiche ad essi relative.

IBM Cloud Pak for Data utilizza efficacemente le seguenti funzionalità per fornire i dati pronti per l'utilizzo aziendale che un'architettura Data Fabric richiede. Tutte queste funzionalità svolgono un ruolo di supporto per l'architettura Data Fabric.

1. Knowledge core basato sui metadati Facilita il rilevamento di origini e cataloghi di dati; arricchisce le risorse di dati ed esegue analisi per estrarre insight per una maggiore automazione utilizzando l'AI. Il knowledge core viene utilizzato per potenziare il marketplace con la ricerca semantica.

- 2. Marketplace di dati self-service** Il catalogo dati di nuova generazione che aiuta i consumatori di dati, come ad esempio gli analisti di business, a recuperare i dati da tutto il panorama dei dati dell'azienda.
- 3. Integrazione intelligente** Consente il consumo di dati tramite l'estrazione, la virtualizzazione, la trasformazione e lo streaming dei dati. Si integra con il knowledge core per automatizzare l'integrazione dei dati e si avvale dell'intelligence per decidere quale approccio di integrazione è più adatto in base ai carichi di lavoro e alle politiche relative ai dati. Si può anche utilizzare per la preparazione dei dati come parte dei carichi di lavoro di ingegneria dei dati o per creare prodotti di dati. Infine, fornisce la possibilità di pubblicare gli aggiornamenti per i prodotti di dati.
- 4. Governance** Cataloga e cura i metadati, definisce le politiche per la privacy relative ai dati, cura i dati, acquisisce la derivazione dei dati ed esegue altre attività relative alla sicurezza e alla conformità. Questo livello riconosce il formato dei dati (ad esempio, strutturati rispetto a non strutturati) e il significato dei dati (ad esempio, pubblici rispetto a protetti) e applica le politiche corrette ad ogni bit di dati e ad ogni potenziale utente. Invece di applicare manualmente standard e regole ai dati, questa funzionalità della piattaforma integrata fa sì che possano essere applicati a livello di organizzazione e propagarsi attraverso le varie risorse di dati, in base alla necessità. I modelli analitici in diversi strumenti possono dialogare tra loro; l'applicazione della politica relativa ai dati a livello granulare può essere ampiamente automatizzata.
- 5. Sviluppo e operazioni unificati** Consente un ciclo di vita unificato per configurare ed eseguire tutti gli aspetti della piattaforma dati in produzione.

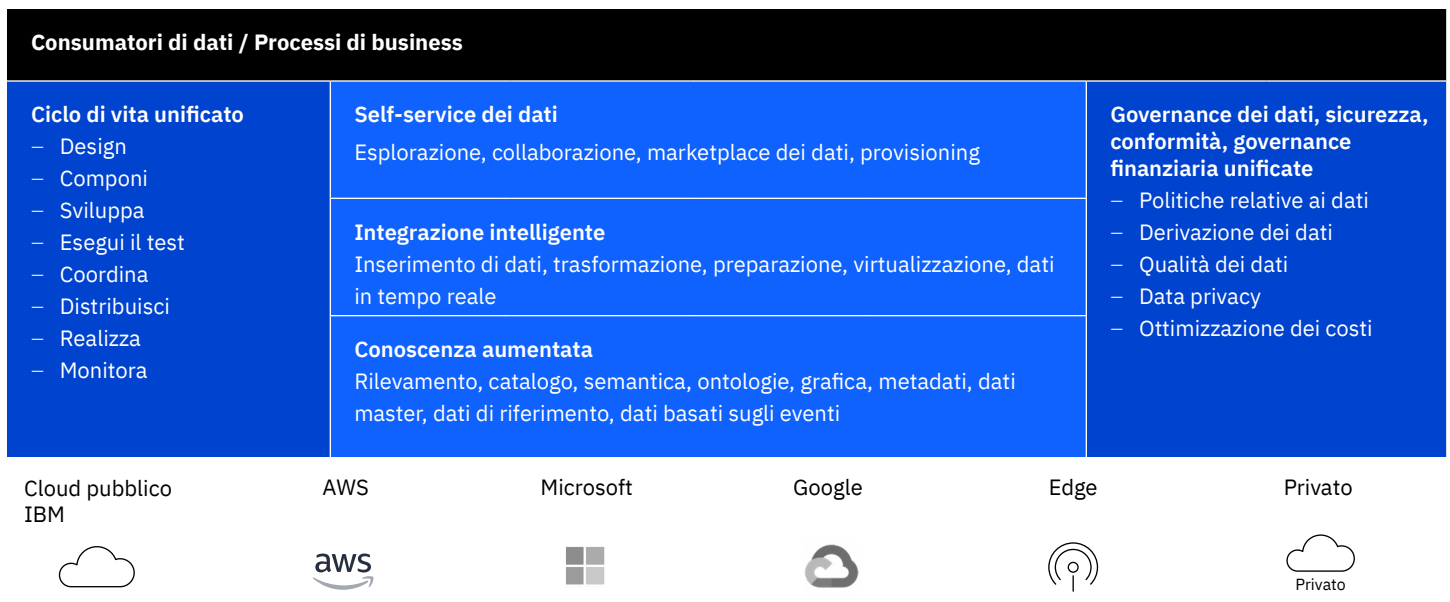


Figura 1. Le funzionalità della piattaforma IBM Cloud Pak for Data, che supportano nel loro insieme lo sviluppo dell'AI e l'architettura Data Fabric

Data Fabric in azione

Il modo migliore per comprendere il valore di un'architettura Data Fabric è quello di esaminare i vantaggi di business che organizzazioni reali hanno raggiunto attraverso l'implementazione di un'architettura Data Fabric su una piattaforma di insight.

Azienda energetica

La sezione riguardante le sfide all'inizio di questo white paper ha introdotto il caso di un'azienda energetica nordamericana in fase di trasformazione digitale.

Questo cliente si è rivolto a IBM che ha implementato un'architettura Data Fabric basata su IBM Cloud Pak for Data.

Con l'implementazione di un'architettura Data Fabric flessibile e integrata, l'azienda è stata in grado di realizzare una serie di importanti progetti di dati in diverse unità, tra cui:

- eMobility
- Rilevamento di documenti relativi alle operazioni del gas, inclusa l'estrazione di quelli scritti a mano
- Segmentazione dei clienti per l'elettricità e previsione del carico
- Gestione degli asset
- Impatti del COVID-19 sul carico
- Modello di rischio in caso di ritorno al lavoro

Con un'architettura Data Fabric, questa organizzazione è in grado di:

- Fornire a più unità di business l'accesso diretto ai dati tramite una piattaforma di insight self-service.
- Utilizzare una funzionalità di calcolo adatta allo scopo per eseguire in modo efficiente modelli su miliardi di righe di dati.
- Sviluppare in modo collaborativo modelli e distribuirli facilmente per fornire insight a tutta l'azienda.

Ridurre gli spostamenti e migliorare la supervisione

Un grande fornitore di servizi industriali ha migliorato la governance dei dati e facilitato la conformità normativa.

Questa organizzazione ha avuto difficoltà nello spostare grandi quantità di dati nel proprio data lake basato sul cloud. Avevano la doppia esigenza di applicare un solido livello di governance a ogni istanza di accesso ai dati e di determinare la qualità dei dati prima di fornire l'accesso agli utenti di business.

Implementando un'architettura Data Fabric basata su IBM Cloud Pak for Data, questa organizzazione ha apportato enormi miglioramenti alla governance dei dati, alla conformità dei dati e ai processi di trasformazione dei dati. Con l'implementazione di un'architettura Data Fabric, possono fornire agli utenti di business un accesso facile e protetto a centinaia di origini dati nel loro data lake basato sul cloud e alle origini dati SAP on-premises.

Stanno anche utilizzando efficacemente gli strumenti acceleratori della normativa di settore per analizzare le origini dati per rilevare possibili dati PII che sarebbero soggetti alla regolamentazione GDPR e CCPA. L'applicazione automatica di contrassegni ai dati aiuta a identificare tutte le origini dati che devono essere gestite per il trattamento delle PII sia on-premises che nel cloud.

Con un accesso semplificato e uno screening più rapido delle PII, gli utenti di business possono estrarre dati rilevanti per ottenere insight importanti senza lunghe attese per l'accesso ai dati o senza correre il rischio di esporre dati protetti.

Con un'architettura Data Fabric, questa organizzazione è in grado di:

- Garantire un'adeguata governance dei dati, utilizzando contemporaneamente in modo efficace i dati di tutta l'azienda.
- Ottenere dati affidabili e ridurre l'entità della preparazione dei dati.
- Fornire assistenza per quanto riguarda la conformità alle normative sulla privacy, quali ad esempio GDPR e CCPA.

Data Fabric è il prossimo passo trasformativo per la tua azienda

Con un'architettura Data Fabric basata sulla tecnologia IBM Cloud Pak for Data, è possibile iper-automatizzare il rilevamento, la governance e l'utilizzo dei dati in un panorama di dati ibrido e multicloud. Impiegare un'architettura Data Fabric per abilitare una produzione di time-to-value più veloce per gli utenti di business, una produttività più elevata per l'ingegneria dei dati e le operazioni e una maggiore governance e fedeltà alla conformità.

[Per iniziare, prova gratuitamente IBM Cloud Pak for Data.](#)

[Pianifica un consulto individuale e gratuito con un rappresentante di IBM Data and AI.](#)



© Copyright IBM Corporation 2021

IBM Italia
Circonvallazione Idroscalo
20090 Segrate (MI)
Italy

Prodotto negli Stati Uniti d'America
Maggio 2021

IBM, il logo IBM e IBM Cloud Pak sono marchi registrati di International Business Machines Corporation negli Stati Uniti e/o in altri paesi. Altri nomi di prodotti e servizi potrebbero essere marchi di IBM o di altre società. Un elenco aggiornato dei marchi IBM è disponibile all'indirizzo ibm.com/trademark.

Questo documento è aggiornato alla data iniziale della pubblicazione e può essere modificato da IBM senza darne preavviso. Non tutte le offerte sono disponibili in ogni paese in cui IBM opera.

I dati relativi alle prestazioni e gli esempi relativi ai clienti, citati nel presente documento, vengono presentati a scopo meramente esplicativo. Le prestazioni reali possono variare a seconda delle specifiche configurazioni e condizioni operative. È responsabilità dell'utente valutare e verificare il funzionamento di qualsiasi altro prodotto o programma con prodotti e programmi IBM. LE INFORMAZIONI CONTENUTE IN QUESTO DOCUMENTO SONO FORNITE "NELLO STATO IN CUI SI TROVANO" SENZA ALCUNA GARANZIA, ESPRESSA O IMPLICITA, SENZA GARANZIE DI COMMERCIALIZZABILITÀ O IDONEITÀ AD UNO SCOPO PARTICOLARE E SENZA ALCUNA GARANZIA O CONDIZIONE DI NON VIOLAZIONE. I prodotti IBM sono garantiti in accordo ai termini e alle condizioni dei contratti che ne regolano la fornitura.

Dichiarazione di procedure di sicurezza valide: la sicurezza dei sistemi IT implica la protezione dei sistemi e delle informazioni attraverso prevenzione, rilevamento e risposta ad accesso improprio dall'interno o dall'esterno dell'azienda. L'accesso improprio può portare all'alterazione, alla distruzione, all'appropriazione abusiva o all'uso non lecito delle informazioni, oppure può portare a danni o all'uso non lecito dei sistemi, che includono l'uso per attacchi ad altri. Nessun sistema o prodotto IT dovrebbe essere considerato completamente sicuro e nessun singolo prodotto, servizio o misura di sicurezza può risultare completamente efficace nel prevenire un uso o un accesso improprio. Sistemi, prodotti e servizi IBM sono progettati per essere parte di un approccio di sicurezza completo, rispettoso della legge, che coinvolgerà necessariamente ulteriori procedure operative e può richiedere altri sistemi, prodotti o servizi per ottenere una maggiore efficacia. IBM NON GARANTISCE CHE SISTEMI, PRODOTTI O SERVIZI SIANO ESENTI DA, O RENDERANNO L'AZIENDA ESENTE DA, LA CONDOTTA MALEVOLA O ILLEGALE DI UNA QUALSIASI PARTE.

Il cliente è responsabile per la garanzia di conformità con i requisiti legali. IBM non fornisce consulenza legale né dichiara o garantisce che i propri servizi o prodotti assicurino che il cliente sia conforme alle normative vigenti.

- 1 Uno studio Forrester New Technology: Projected Total Economic Impact™ commissionato da IBM, febbraio 2020
- 2 Le misurazioni delle prestazioni sono state raccolte all'interno di un ambiente di test controllato presso gli IBM Silicon Valley Labs utilizzando la virtualizzazione dei dati IBM rispetto a varie origini dati da 100TB. Le misurazioni effettuate nel maggio 2019 e i guadagni di prestazioni sono paragonati alla federazione IBM.
- 3 Basato sull'esperienza di un cliente IBM.