



Il percorso verso AI e dati pronti per il business inizia con l'architettura delle informazioni

Scopri un ambiente di analytics affidabile incentrato su governance e catalogazione

Introduzione

Indice

- Introduzione
- Gli elementi indispensabili per una analytics affidabile
- Il machine learning accelera la governance
- Un'unica base per più finalità
- Conclusioni

Consigli chiave

- Le esigenze di conformità spingono le aziende a implementare strategie di governance dei dati continue e vantaggiose
- Il machine learning (ML) automatizza le iniziative per la governance e l'integrazione su vasta scala e permette di superare le sfide presentate dai volumi di dati di grandi dimensioni e dai limiti delle capacità umane
- La governance dei dati è utile sia in ambienti on-premises che multicloud.

I dati sono sempre più numerosi e vari nelle aziende di tutti i tipi. Negli ambienti multicloud, la vasta gamma di fonti di dati aumenta in maniera esponenziale il flusso di informazioni in arrivo, dall'IoT (Internet of Things) e i social media, ai dispositivi mobili, alle implementazioni della realtà virtuale e la traccia ottica. Benché oggi le organizzazioni stiano investendo prontamente nell'AI (artificial intelligence), la maggior parte di esse non comprende in modo adeguato i dati o non acquisisce dati di qualità tale da trarre vantaggio dalle soluzioni AI. In molte aziende i dati sono inaccessibili, inaffidabili o non conformi alle regole di protezione e privacy.

Le normative globali, come il GDPR (General Data Protection Regulation), il CCPA (California Consumer Privacy Act) e l'LPGD (Brazil's Lei Geral de Proteção de Dados) puntano l'attenzione sui dati personali dei clienti e dei dipendenti. Tali normative offrono alle aziende un'opportunità per trasformare e creare nuovi modelli di business basati sui dati, nonostante le penali che possono derivare dalla non conformità, cosa che può rallentare la produttività o danneggiare il valore del brand. Per rispettare le normative sulla privacy e proteggere le informazioni personali, le aziende devono prima scoprire e classificare i loro diversi tipi di dati. Le aziende che faticano a raccogliere o a utilizzare in maniera appropriata i dati dei clienti possono incontrare seri problemi. Per superare tali sfide, le aziende stanno implementando [architetture di informazioni regolamentate](#) che rispettino le normative pur continuando a fornire prestazioni e innovazione basata sui dati.

Gli elementi fondanti per una analytics affidabile

Affrontare il problema delle normative sulla privacy dei dati come un obbligo e un'opportunità per modernizzare le infrastrutture di dati, offre un notevole vantaggio. Ciò incoraggia le organizzazioni a implementare strategie di governance dei dati che generino nuovi modelli di business e consentano di ottenere utili insight. Le iniziative UGI (Unified governance and integration) si applicano ai dati, strutturati e non, in cloud pubblici e privati. L'implementazione dell'UGI per raggiungere la conformità è già di per sé di significativa importanza, ma influisce anche su altre aree di un'azienda, in particolare sulla governance dei modelli AI per i data scientist.

Quando un'azienda utilizza la governance dei dati per acquisire dati corretti, gli utenti sanno che i dati provengono da source affidabili. Sanno come vengono utilizzati i dati nell'azienda e che ciò migliorerà qualsiasi progetto di analytics. Le iniziative di analytics necessitano che i dati affidabili funzionino in modo efficace, a prescindere da quanto avanzati siano gli strumenti. I vantaggi della possibilità di utilizzare dati attendibili e pronti per il business sono illimitati. L'analytics può suggerire la progettazione di nuovi prodotti e programmi di marketing e migliorare le vendite, la supply-chain o iniziative di assistenza al cliente. L'analytics può anche rilevare inefficienze operative che, una volta eliminate, consentono di migliorare l'agilità e accrescere i guadagni.

L'implementazione della governance di dati e AI nella tua organizzazione coinvolge i seguenti sette punti cardine.

Rilevamento e qualità dei dati

Le aziende possono non essere consapevoli dell'enorme quantità di dati memorizzati in seguito alle loro attività commerciali. Il primo passo nella governance dei dati consiste nell'inventariare i dati dell'organizzazione. Inizia concentrandoti sui dati in uno specifico progetto e successivamente includi altri casi di business per coprire più settori dell'azienda. Memorizzare e gestire dati ridondanti, obsoleti o banali (detti ROT) non solo è dispendioso ma [crea confusione durante il processo decisionale e le operazioni](#). Può anche rendere più difficile rispettare la conformità e quindi vanificare gli sforzi nell'analytics. I dati devono rispettare e garantire nel tempo determinati livelli di qualità affinché l'utilizzo nel downstream dia risultati positivi.

Catalogazione

Una volta rilevati e profilati i dati, vanno catalogati utilizzando tag di metadati per identificarne il tipo, l'uso, la proprietà, la derivazione e altro ancora. Poiché le aziende in alcuni settori condividono esigenze comuni, dei modelli di settore precostruiti possono accelerare il processo di catalogazione utilizzando tassonomie e termini di business già pronti. Con i miglioramenti nel machine learning, i termini di business possono essere automaticamente associati per creare un [catalogo aziendale](#) in poche ore. Le basi di catalogazione che l'UGI fornisce, permettono alle organizzazioni di gestire i modelli AI, i notebook e altre fonti di dati, creando una libreria centralizzata delle conoscenze dell'organizzazione. Tali basi sono una risorsa per molti utenti di dati nell'organizzazione, tra cui i data engineer, i data steward e gli utenti della LOB (line-of-business) come gli analisti, i data scientist e i responsabile del marketing.

Spostamento, trasformazione e sincronizzazione dei dati

I dati da più origini possono essere facilmente integrati, trasformati e condivisi con altri sistemi come necessario, sia fisicamente che virtualmente. Questo processo riunisce i dati strutturati e quelli non strutturati e consente l'integrazione con tecnologie open come Apache Atlas e Hadoop. La creazione di flussi di dati e sincronizzazione automatizzati garantisce che nei data lake, nei data warehouse, nei data mart e nelle soluzioni punto d'impatto siano sempre disponibili i dati più recenti. Man mano che le quantità di dati aumentano, la replica supporta ampi volumi con bassa latenza. Le aziende possono utilizzare la virtualizzazione senza dover spostare i dati, in base alle proprie esigenze.

Entro il 2019, la produzione di analytics degli utenti di business con capacità self-service supererà quella dei data scientist professionisti.

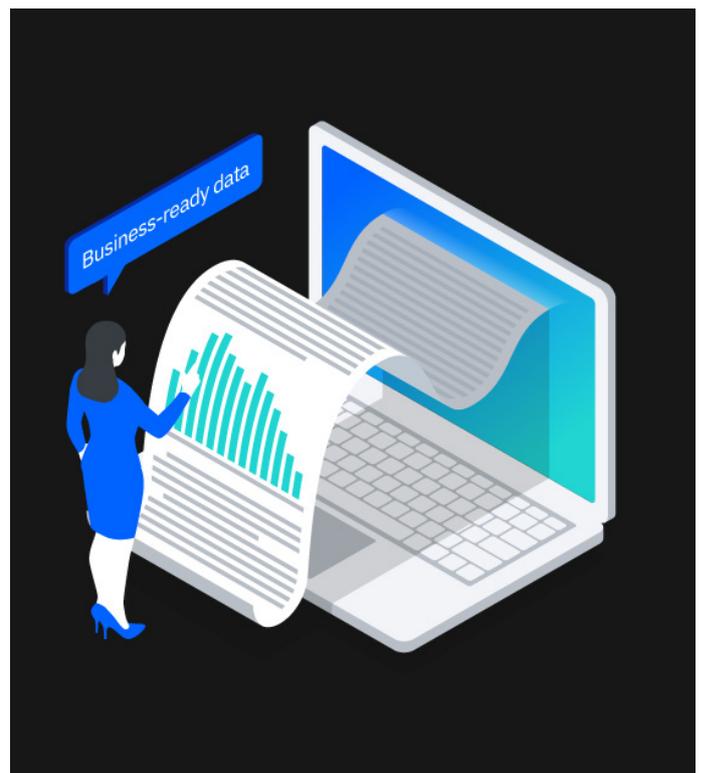
Master data management

Secondo lo studio di Gartner, entro il 2019, la produzione di analytics degli utenti di business con capacità self-service supererà quella dei data scientist professionisti. È essenziale per le organizzazioni fare affidamento su una visione completa, affidabile e unificata delle entità critiche come i clienti, i prodotti e gli account. [Le moderne implementazioni MDM \(master data management\)](#) forniscono esplorazione analitica basata su grafici, un motore di corrispondenza estremamente preciso, un approccio iniziale ai dati basato sulla selezione di algoritmi di corrispondenza e processi di gestione regolati dal machine learning. Inoltre, le soluzioni MDM forniscono agile accesso self-service, strumenti di governance e funzioni dashboard intuitive.

Privacy e protezione dei dati

Le aziende devono proteggere e salvaguardare in maniera proattiva i loro asset di informazioni strategici e sensibili. La gestione del ciclo di vita dei dati va dalla creazione allo smaltimento, e si avvale di prassi come la gestione di record, controversie e storage degli archivi. Grazie al cognitive learning applicato ai documenti attuali e storici di un'organizzazione, i rischi possono essere automaticamente identificati in base al contesto dell'organizzazione stessa.

I dati governati, sia per operazioni di business che per scopi di conformità, sono pronti per il business e possono essere immediatamente utilizzati in qualsiasi processo decisionale, sia di miglioramento che di innovazione. Man mano che le quantità di dati aumentano, la replica supporta ampi volumi con bassa latenza. Le aziende possono utilizzare la virtualizzazione senza dover spostare i dati, in base alle proprie esigenze.



Il machine learning accelera la governance

Il machine learning oggi potenzia l'intelligenza umana e supplisce alle significative limitazioni delle capacità umane, grazie alle recenti scoperte tecnologiche. Automatizza le iniziative per la governance e l'integrazione su vasta scala e permette di superare le sfide presentate dai volumi di dati di grandi dimensioni, portando a una corretta governance dei dati in tutta l'azienda. Ad esempio, se un'azienda dispone di 20.000 termini di dati, generalmente

impiega sei mesi e un team di sei persone per classificare manualmente i termini per ottenere l'analytics in modo preciso e affidabile. Con il machine learning, lo stesso processo può essere completato in pochi giorni, o persino ore, a seconda della quantità di asset di dati. Una tale capacità di accelerazione riduce drasticamente i costi del processo di governance. Il machine learning può rendere più gestibili gli obblighi di conformità e preparare la strada a iniziative di analytics che siano realmente produttive.

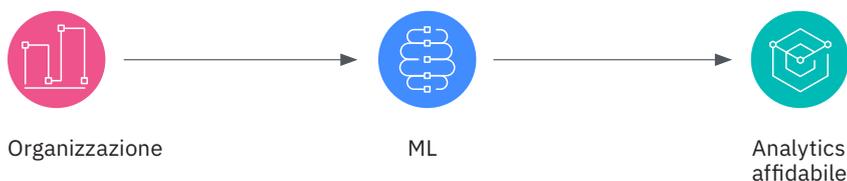
Classificazione manuale

Sei mesi per il completamento



Classificazione ML

Giorni o ore per il completamento



Un'unica base per più finalità

Quando la governance ha solide basi, può essere utilizzata in più unità di business e in tutta l'azienda, come descritto nei seguenti esempi di uso comune.

Data lake governato

Un ampio numero di aziende in una vasta gamma di settori d'industria ha avviato progetti di predictive analytics di big data. Il primo fondamentale passo consiste nel depositare grandi quantità di dati, strutturati e non, in un data lake.

Le organizzazioni hanno utilizzato soluzioni Hadoop o di codifica manuale nei propri data lake. L'assenza di una struttura di governance associata alla mancanza di strategia nella gestione di standard di dati, termini di business, provenienza, uso e qualità possono portare al data swamp. Ciò si verifica quando gli utenti non conoscono e non considerano affidabili i propri dati. Le aziende stanno imparando che per ottenere risultati positivi i loro data lake devono avere una strategia e una governance.

In una [implementazione governata](#), i dati in un data lake vengono associati a termini di business semplici da comprendere per qualsiasi utente di dati e congruenti in tutta l'organizzazione. La possibilità degli utenti di avere tali dati, accelera il time-to-value praticamente per qualsiasi progetto di data science, esplorazione di dati o AI self-service, contribuendo a fornire agilità. Questa capacità pone le basi necessarie per supportare ambienti multicloud, architetture on-premises e una varietà di origini di dati.

Modernizzazione dell'applicazione

Le aziende stanno effettuando grossi investimenti nella modernizzazione delle applicazioni per aumentare l'efficienza, ridurre i costi e ottenere vantaggi competitivi. A seconda dell'azienda, le idee sulla modernizzazione delle app possono manifestarsi in vari modi. Le considerazioni principali includono i test della gestione dei dati, della virtualizzazione dei dati e della connettività. Le aziende oggi utilizzano metodologie

agili per testare e sviluppare simultaneamente con [l'accesso virtuale ai dati](#). Fanno anche affidamento sulla connessione di applicazioni di business e dati tramite funzioni di integrazione flessibili.

Una visione a 360 gradi dei dati del cliente

È importante che i dipendenti abbiano [informazioni master in una singola vista attendibili, aggiornate e precise](#) relativamente ai clienti, ai prodotti e ad altre entità. Informazioni errate o non aggiornate possono danneggiare le interazioni con i clienti e diminuire la fiducia, portando a una riduzione dei profitti o a un aumento dei costi della supply chain. I dati che originano dai processi UGI aiutano le interazioni con i clienti aumentando la fiducia, la lealtà al brand e migliorando l'agilità della supply chain.

Ottimizzazione del data warehouse aziendale

Con l'EDW (enterprise data warehouse), l'ottimizzazione dell'architettura rappresenta un aggiornamento e un determinante cambiamento nel modo di accedere ai dati, memorizzarli, prepararli, governarli e analizzarli. Uno dei più [efficaci approcci all'ottimizzazione](#) consiste nell'offload di lavori ETL (extract, transform, load), di dati non più utilizzati e di dati richiesti in modelli di esplorazione. Questo processo non solo riduce i costi, ma permette di combinare i dati con altri tipi di dati in ambienti come data lake governati per un'esplorazione dei dati dinamica.

Conformità normativa

Una base di analytics affidabile consente e [accelera la conformità](#) quando si tratta di obblighi normativi. La cosa più importante è che il percorso verso la protezione dei dati personali inizia con il riconoscimento di quali siano i dati personali, in modo che un'azienda riesca a capire di quali dati personali dispone. Il catalogo dei dati di base contiene regole di governance per ottenere qualità, quantità e analisi di dati, oltre alle policy per la conformità.

Conclusioni

Man mano che le aziende affrontano il percorso verso la trasformazione digitale, i business leader diventano consapevoli dei vantaggi della governance nei dati e nei modelli AI, sia in ambienti on-premises che multicloud. Concentrandosi su solide procedure di governance, le organizzazioni preparano i dati e l'AI non solo per l'elaborazione analitica e gli insight, ma anche per la conformità alle normative che devono rispettare. Quando i volumi di dati sono enormi, le funzionalità ML e AI potenziano l'intelligenza umana in attività come l'associazione dei dati, la catalogazione, la corrispondenza con volumi di grandi dimensioni e il raggiungimento della qualità.

I business leader lungimiranti sanno che il tempo impiegato nella creazione di solide basi per l'UGI verrà ripagato con significativi dividendi oggi e nel prossimo futuro. Sono consapevoli che le loro aziende otterranno importanti vantaggi dall'adozione della governance come strumento per raggiungere l'ottimizzazione del business, l'innovazione e la conformità in tutte le iniziative sui dati e sull'AI. È cruciale avvalersi di soluzioni che contemplino la gestione delle operazioni sui dati, dalla creazione al consumo. La semplificazione di queste operazioni richiede economie di scopo, scala e condivisione.

Scopri di più

Le soluzioni IBM UGI ti permettono di creare una solida base di analytics affidabile per promuovere l'AI in scala.

[Visita il sito web](#)
ibm.com/unified-governance-integration

[Scopri di più](#)
Continua ad acquisire informazioni sulla governance dei dati cognitivi e sulle soluzioni all'avanguardia di IBM per l'ML e l'AI. [Leggi il whitepaper.](#)

[Parla con un esperto](#)
Mettiti in contatto con uno dei nostri esperti UGI che hanno lavorato con migliaia di clienti per creare strategie vincenti di AI, analytics e dati.
[Prenota una consulenza di 30 minuti](#)

IBM Italia S.p.A.
Circonvallazione Idroscalo
20090 Segrate (Milano)
Italia

La home page di IBM Italia si trova all'indirizzo:
ibm.com

IBM, il logo IBM e ibm.com sono marchi di International Business Machines Corp., registrati in diverse giurisdizioni nel mondo. Altri nomi di prodotti o servizi possono essere marchi di IBM o di altre società. Un elenco aggiornato dei marchi IBM è disponibile sul web nella pagina "Informazioni su copyright e marchi" all'indirizzo ibm.com/legal/copytrade.shtml

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

Gli esempi di clienti citati sono presentati solo a scopo illustrativo. Le prestazioni reali possono variare a seconda delle specifiche configurazioni e condizioni operative.

LE INFORMAZIONI CONTENUTE IN QUESTO DOCUMENTO SONO FORNITE NELLO STATO IN CUI SI TROVANO, SENZA ALCUNA GARANZIA, ESPRESSA O IMPLICITA, INCLUSE, A TITOLO DI ESEMPIO, GARANZIE IMPLICITE DI COMMERCIALIZZABILITÀ E DI IDONEITÀ PER UNO SCOPO SPECIFICO E DI NON VIOLAZIONE. I prodotti IBM sono garantiti secondo i termini e le condizioni dei contratti che ne regolano la fornitura.

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.

© Copyright IBM Corporation 2019