



INDUSTRY 4.0 for VET

3. BIG DATA

3.1 INTRODUZIONE

La prima introduzione

Riesci a immaginare che un essere umano impiegherebbe circa 181 milioni di anni per scaricare tutti i dati disponibili su Internet? Queste grandi quantità di dati che sono disponibili oggi e il modo in cui vengono elaborati, fa parte dell'ambito dei **Big Data**.

Come vedrai in questa unità, ci troviamo di fronte a questo discorso quasi ogni giorno, spesso a nostra insaputa. Imparerai a conoscere i **vantaggi** ma anche i **pericoli** dei Big Data e perché la **corretta gestione** di queste grandi quantità di dati è spesso più importante dei dati stessi.



La rilevanza pratica – per queste ragioni avrai bisogno di queste abilità

Non solo gli specialisti IT, ma quasi tutti incontrano big data in situazioni quotidiane, come quando si visita un medico, si naviga su social media come Facebook e Instagram, si compie una ricerca su Google o su una piattaforma connessa alla rete

Sapere come vengono utilizzate grandi quantità di dati e quali **opportunità** ma anche **pericoli** siano associate può essere rilevante sia per il tuo uso personale di Internet che per la tua vita professionale, magari al servizio di un'azienda che analizza grandi quantità di dati.

Obiettivi dell'apprendimento e competenze

Questa unità di apprendimento ti offre una conoscenza di base dei Big Data. Conoscerai il **modello 3-V** e imparerai come vengono raccolte e analizzate grandi quantità di dati. Imparerai quindi gli **scopi** per i quali vengono utilizzate le **conoscenze** acquisite da grandi quantità di dati e i **rischi** connessi alla loro gestione. Scoprirai perché la protezione dei dati è diventata sempre più importante negli ultimi anni e capirai che la gestione dei Big Data pone grandi sfide sia per le aziende che per i privati.

Obiettivi dell'apprendimento

Capire e descrivere il termine Big Data

Sapere come usare i Big Data

Capire e spiegare come grandi quantità di dati vengano gestiti e utilizzati

Sapere quali rischi e opportunità siano associati ai Big Data

3.2 Cosa sono I Big Data?

Sapevi che circa il 90% di tutti i dati disponibili oggi nel mondo è stato generato negli ultimi anni? Grazie alle nuove tecnologie di informazione e comunicazione, il **volume di dati** in tutto il mondo è cresciuto incredibilmente e offre possibilità precedentemente sconosciute. Il termine **Big Data** sta per questo volume di dati strutturati e non strutturati, che non possono essere elaborati con software o hardware convenzionali a causa delle loro dimensioni.



Questi volumi di dati vengono creati, tra le altre cose, con ciascuno dei nostri clic su Internet. Questo può essere, ad esempio, un acquisto su Amazon, una chiave di ricerca su Google, le attività compiute sui social network come Instagram o Facebook, ecc.

Tuttavia, grandi quantità di dati da sole non sono sufficienti a produrre Big Data. Solo l'analisi e l'elaborazione di questi volumi di dati, ad es. da un'azienda, produce i Big Data. Nel 2001, l'analista Doug Lane ha creato una definizione di Big Data con il suo **modello a 3 V** ancora oggi riconosciuto. Secondo Lane, i Big Data hanno le seguenti tre caratteristiche:

- **Volume:** Le aziende raccolgono grandi volumi di dati da varie fonti. Questi includono dispositivi intelligenti (IoT) come telefoni cellulari, video, social media, ecc. In passato non sarebbe stato possibile archiviare questo grande volumi di dati; oggi esistono piattaforme di archiviazione per questo scopo.
- **Velocità:** Le aziende sono attualmente invase da flussi di dati a velocità senza precedenti che devono essere elaborate rapidamente.
- **Varietà:** I dati raccolti sono eterogenei e hanno una grande varietà di formati: i dati numerici, disponibili in forma strutturata e archiviati in normali banche dati, possono far parte di Big Data, nonché documenti di testo non strutturati, creati da transazioni finanziarie o e-mail.

Definizione
<p>Big Data ... significa una grande quantità di dati disponibili che vengono analizzati ed elaborati per uno scopo specifico. Secondo Doug Lane, i Big Data sono caratterizzati da volume, velocità e varietà.</p>

Big Data vs. Small Data



“Let’s shrink Big Data into Small Data ...
and hope it magically becomes Great Data.”

Diversamente dai Big Data, Small Data si riferisce ai dati in un volume e in un formato accessibili all'uomo. I seguenti punti mostrano come i Big Data possono essere distinti dai Small Data:

- **Obiettivi:** gli Small Data vengono utilizzati per un obiettivo definito, l'uso dei Big Data si sviluppa spesso in modo imprevisto.
- **Luogo:** I piccoli dati vengono generalmente archiviati in un unico posto, di solito in un file sul PC, mentre i big data vengono generalmente distribuiti su numerosi file su server diversi situati in paesi diversi.
- **Struttura dei dati:** I piccoli dati sono strutturati in linea retta, mentre i big data possono essere non strutturati e possono avere molti formati di file da campi diversi.
- **Preparazione dei dati:** di solito solo un utente finale è coinvolto nella preparazione di piccoli dati. Nel caso dei Big Data, tuttavia, spesso accade che un gruppo di persone prepari i dati, un altro gruppo analizzi i dati e infine un terzo gruppo utilizzi i dati. Ognuno di questi gruppi può avere obiettivi diversi.
- **Durata:** Gli Small data vengono generalmente conservati per un certo periodo di tempo dopo il completamento di un progetto. Nel caso dei Big Data, tuttavia, i dati rimangono memorizzati per un periodo di tempo illimitato.
- **Origine:** gli Small Data vengono memorizzati in breve tempo e in unità di misura specifiche. I Big Data, d'altra parte, provengono da diversi luoghi, paesi, aziende, organizzazioni, ecc.
- **Riproducibilità:** Small Data possono generalmente essere completamente riprodotti. I Big Data, al contrario, provengono da molte fonti diverse e sono disponibili in talmente tante forme che la loro riproduzione è impossibile

- **Qualità:** i significati dei dati in un set di small data non sono ambigui, pertanto questi dati possono essere descritti. I Big Data, al contrario, sono molto più complessi e possono anche contenere informazioni non identificabili che non hanno un significato specifico. Ciò può ridurre la qualità delle informazioni
- **Analisi:** un singolo processo è in genere sufficiente per l'analisi di piccoli dati, poiché i dati vengono analizzati da un solo file di computer. Nel caso dei Big Data, i dati devono essere estratti, controllati, ridotti, ecc. in un processo che richiede tempo.

Come puoi vedere dalla distinzione tra Big Data e Small Data, i Big Data sono letteralmente spesso difficili da comprendere.

3.3 Possibili usi e opportunità dei Big Data

L'analisi di grandi quantità di dati consente di ottenere approfondimenti. Questi risultati possono servire come base per le decisioni, ad esempio, in merito alla direzione strategica dell'azienda. Le aziende, ad esempio, vogliono saperne di più sulle preferenze dei propri clienti al fine di adattare la loro gamma di prodotti, pubblicità e così via.

Il **Deep Learning** utilizza anche Big Data: questo è un metodo speciale di elaborazione delle informazioni e una branca del **machine learning**. Una macchina viene "alimentata" con grandi quantità di dati, che vengono analizzati e utilizzati per addestrare la macchina. La macchina è in grado di collegare tra loro nuove informazioni e su questa base può fare previsioni e prendere le proprie decisioni. Tuttavia, il risultato è buono solo come i dati, da cui la macchina ha "imparato"



Un esempio è un sistema di traduzione automatica che "impara" a tradurre correttamente i termini tecnici in un'azienda inserendo dati (traduzioni esistenti).

Inoltre, le autorità e i servizi segreti utilizzano grandi quantità di dati per rilevare discrepanze e anomalie che potrebbero indicare attività criminali o terroristiche. Nella scienza, grandi quantità di dati vengono utilizzati per studiare complessi fenomeni naturali come i cambiamenti climatici o il verificarsi di terremoti ed epidemie.

Tuttavia, le grandi quantità di dati non sono sempre gestite in modo responsabile. Alcune società o istituzioni non aderiscono alle normative sulla protezione dei dati, il che significa che le informazioni vengono rese pubbliche. Questo può essere banale, ma in alcuni casi può anche essere pericoloso e portare a frodi e ricatti.

Esempio
Nel 2015, il portale di incontri di Ashley Madison, in cui le persone in cerca di un'avventura extraconiugale possono creare un profilo, è diventato vittima di un attacco di un hacker. Di conseguenza, le informazioni sulle persone registrate sul portale sono diventate disponibili su Internet. Le informazioni sulle avventure delle celebrità e le informazioni personali come i numeri delle carte di credito sono diventate pubbliche. Inoltre, le persone interessate sono state invitate via e-mail a pagare un riscatto in modo che il loro compagno di vita non venisse a conoscenza del profilo sul portale di lancio.
Ricorda
Grandi quantità di dati possono essere utilizzate per i seguenti scopi, tra gli altri: <ul style="list-style-type: none"> • Orientamento strategico di un'azienda • Deep Learning • Lottare contro il crimine e il terrorismo • indagine scientifica sui fenomeni naturali (ad es. Terremoti e cambiamenti climatici) • valutazioni illegali che possono comportare ricatti o frodi <p>Il fattore decisivo per quanto riguarda i Big Data non sono tanto i dati stessi quanto ciò che accade ad essi.</p>

Le aziende in particolare beneficiano dell'analisi e della valutazione dei Big Data. Sia consciamente che inconsciamente, oggi generano e archiviano enormi quantità di dati. Di seguito, imparerai in dettaglio quali possibilità offre alle aziende la corretta analisi di grandi quantità di dati.

Prendere decisioni

Analizzando le grandi quantità di dati generati nell'azienda, le aziende possono identificare i modelli e filtrare le informazioni. Ciò consente alle aziende di prendere decisioni aziendali migliori che aumentano il successo dell'azienda. Valutando i dati della macchina, per Esempio, è possibile calcolare a quali intervalli si rompe una macchina. L'azienda può utilizzare questa conoscenza per riparare la macchina prima che si guasta. I big data vengono anche utilizzati nel settore finanziario e assicurativo per calcolare meglio i rischi.

Esempio
La signora Rossi ha 47 anni e vorrebbe stipulare un'assicurazione sanitaria privata. Quando visita il suo broker assicurativo, è sorpresa per gli alti costi e le richieste. Si scopre che il suo fornitore analizza grandi quantità di dati al fine di calcolare meglio i costi assicurativi dei singoli clienti. L'azienda scopre, per esempio, quali sono i rischi particolari per la salute delle donne di questa età che, come la signora Rossi, sono fumatori, non hanno figli e non fanno mai sport.

Incremento nell'efficienza

La competitività è molto importante per le aziende. Per stare al passo con la concorrenza, le aziende devono progettare strategie per risparmiare sui costi senza compromettere le prestazioni. Analizzare e collegare grandi quantità di dati aiuta a fare questo.

Esempio
Hai mai sentito dire che i driver UPS girano quasi sempre a destra? Questo perché UPS ha scoperto, attraverso l'analisi dei big data, che questo può fargli risparmiare circa \$ 10 milioni all'anno. Probabilmente ti starai chiedendo come sia possibile: la fusione di vari set di dati, come le statistiche sugli incidenti, i dati sul consumo di carburante, ecc., Ha dimostrato che i veicoli UPS hanno

molte meno probabilità di essere coinvolti in incidenti se non girano a sinistra. Ciò può far risparmiare molto denaro, anche se di conseguenza le rotte diventano più complicate.

Previsione in ricerca e sviluppo

Rendendo gli acquirenti esistenti o potenziali o i clienti consapevoli delle loro preferenze per determinati prodotti, la ricerca può identificare e prevedere le tendenze, progettare strategie di marketing appropriate e sviluppare prodotti su misura. Con i metodi analitici appropriati, è anche possibile, per esempio, prevedere la sicurezza della rottura di un prodotto mentre è ancora in fase di sviluppo.

Esempio

Un operatore di un negozio online installa cookie e tracciamento online e tiene traccia dei movimenti dei suoi visitatori. Può determinare la provenienza dei visitatori, i prodotti su cui fanno clic, la frequenza con cui visitano il sito, ecc. Con l'aiuto di questi dati, l'operatore può adattare i contenuti del sito e i prodotti offerti alle preferenze dei visitatori e quindi aumentare il suo fatturato.

Servizio clienti personalizzato

Memorizzando le decisioni dei clienti, le aziende sono in grado di fornire loro un servizio clienti personalizzato. Ad esempio, se un utente guarda un determinato film o serie su Netflix, il sistema lo salverà e al successivo accesso dell'utente, i consigli si baseranno sui film o sulle serie che l'utente ha già visto. Ma questa offerta personalizzata non sempre soddisfa l'approvazione:

Esempio

Quando il signor Mori si rende conto che i suoi vecchi scarponi da montagna non sono più utilizzabili, cerca su Google "scarponi da montagna nuovi per uomo". È sopraffatto da molte offerte diverse e il signor Mori scopre anche che molti prodotti non possono essere consegnati nel suo paese d'origine, l'Austria. Il signor Mori decide di ottenere una consulenza personale in un negozio specializzato e acquista anche un paio di scarponi da montagna. Tuttavia, vede sempre più pubblicità per gli scarponi da montagna su Internet nei prossimi giorni e settimane, poiché la sua ricerca è stata salvata e analizzata su Google. Mori è irritato e si sente osservato. Decide di non effettuare più ricerche di prodotti su Google in futuro.



Ricapitolando:

Ricorda

Le aziende hanno numerose opportunità di utilizzare i Big Data per avere più successo. Questi includono:

- **Prendere decisioni:**
L'analisi dei Big Data consente alle aziende di prendere decisioni aziendali migliori e valutare meglio i rischi.

- **Incremento nell'efficienza**
L'analisi e il collegamento dei dati (come i dati meteorologici e di congestione con i prezzi del carburante) aiuta le aziende a rendere i processi più efficienti.
- **Previsione in ricerca e sviluppo**
Con l'aiuto dei Big Data, è possibile fare previsioni su tendenze, caratteristiche di un prodotto, ecc.
- **Servizio clienti personalizzato**
Memorizzando le decisioni prese dai clienti, le aziende possono offrire loro un servizio clienti personalizzato alla loro prossima visita.

3.4 Come vengono analizzati i Big Data?

Hai imparato come sono definiti i Big Data e quali opzioni ci sono per l'utilizzo di grandi quantità di dati. In questo capitolo, entreremo nel dettaglio e tratteremo dell'analisi dei Big Data. Questo campo specialistico è noto come **Big Data Analytics**.



Big Data Analytics – Teoria

Il primo passo è quello di raccogliere grandi quantità di dati da diverse fonti, che hanno formati diversi. Questo viene spesso fatto utilizzando lo strumento di ricerca. In seguito i dati vengono preparati per ulteriori elaborazioni. Un problema è spesso che grandi quantità di dati sono disponibili in una forma non strutturata e in formati completamente diversi e quindi non possono essere acquisiti dal software di database convenzionale.

La Big Data Analytics utilizza quindi processi complicati per estrarre e acquisire i dati. I dati vengono quindi analizzati utilizzando uno speciale software per la gestione di Big Data. Infine, i risultati vengono elaborati e presentati.

È importante che il software utilizzato sia in grado di implementare rapidamente molte richieste di ricerca e di importare ed elaborare rapidamente i vari set di dati. Per essere ancora più potenti, molti sistemi non utilizzano lo spazio sul disco rigido (come le applicazioni di database convenzionali) per l'elaborazione dei dati, ma piuttosto la memoria principale, generalmente molto più veloce. In questo modo, la velocità di accesso può essere aumentata e le analisi possono essere eseguite quasi in tempo reale.

Ricorda

L'analisi dei Big Data può essere approssimativamente suddivisa in tre diverse aree:

- Acquisizione di dati da molte e varie fonti utilizzando query di ricerca
- Valutazione e ottimizzazione dei dati raccolti
- Analisi dei dati e sintesi e presentazione dei risultati

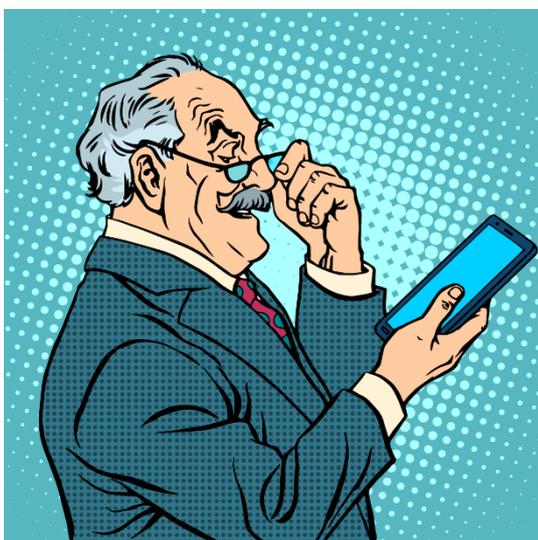
Un software potente e adatto è molto importante per questo.

Big Data Analytics – In pratica

È interessante notare che Big Data Analytics è ancora agli inizi nella maggior parte delle aziende e che le opportunità offerte sono lungi dall'essere esaurite. In media, le aziende analizzano solo poco più di un terzo dei dati generati dal contatto digitale con i propri clienti (ad es. Tramite negozi online o siti Web).

La ragione di ciò sono spesso le rigide normative sulla protezione dei dati che rendono più difficile l'analisi dei Big Data. Le leggi e i regolamenti che regolano la protezione dei dati sono discussi più dettagliatamente nel capitolo seguente. In realtà, tuttavia, per molti aspetti le aziende non sono ancora pronte per utilizzare efficacemente le grandi quantità di dati per se stesse. Le seguenti aree svolgono un ruolo importante:

Prima di tutto, è consigliabile distribuire correttamente i risultati: le fonti di dati dovrebbero provenire da aree diverse, i risultati dovrebbero essere utilizzati in diverse aree dell'azienda. È inoltre necessaria una strategia adeguata: un'azienda dovrebbe sapere a quale scopo vengono analizzate le grandi quantità di dati. Una cultura aziendale adeguata è anche molto importante, le nuove tecnologie, per Esempio, non dovrebbero essere respinte in linea di principio, ma piuttosto considerate realisticamente.



La maggior parte delle aziende non ha un proprio reparto per l'analisi dei dati. Tuttavia, alcuni dipendenti dovrebbero portare con sé le competenze necessarie o acquisirle nei corsi di formazione. Potrebbe essere necessario assumere nuovi dipendenti. Le responsabilità e le autorizzazioni devono essere definite anche all'interno dell'azienda.

Per l'analisi è necessaria una tecnologia efficiente, sotto forma di adeguati strumenti di analisi dei Big Data. Tuttavia, per capire quali strumenti siano adatti occorre riflettere sulla strategia precedentemente definita o sullo scopo dell'analisi. Infine, è essenziale anche un'adeguata strategia di protezione dei dati per garantire che i dati personali degli individui non siano divulgati al pubblico. Un esperto dedicato alla protezione dei dati all'interno dell'azienda assicura che l'analisi dei dati sia conforme alle leggi e ai regolamenti applicabili.

Ricorda

In breve, per far sì che la **Big Data Analytics** possa essere effettuata con successo in una compagnia, occorre:

- Una strategia sui Big Data – definire gli scopi dell'analisi
- Un'adeguata cultura aziendale – apertura alle nuove tecnologie
- Personale con il *know how* necessario – formare o reclutare
- Una tecnologia potente – Strumenti adeguati all'analisi
- Un'appropriata politica di protezione dei dati – in linea con le leggi e i regolamenti

3.5 Sfide e rischi dei Big Data

Nei capitoli precedenti hai assistito alla complessità dell'analisi e dell'utilizzo dei Big Data. Almeno altrettanto complesse sono le sfide e i rischi associati a grandi quantità di dati. Probabilmente la più grande sfida per le aziende in relazione ai Big Data è la protezione dei dati:



Sebbene le aziende abbiano prestato maggiore attenzione alla protezione dei dati negli ultimi anni, ci sono ancora problemi. Ad esempio, i dati personali degli utenti di Internet possono essere utilizzati senza il loro consenso e le persone interessate possono essere identificate, controllate e nel peggiore dei casi ricattate.

Definizione

Dati personali

fa riferimento a dati che si riferiscono a una persona e consente di trarre conclusioni sulla sua personalità. Ciò include, ad esempio, il numero di targa di Mario Bianchi, la data di nascita del tuo vicino o il saldo del conto di Bill Gates.

Un esempio di violazione della protezione dei dati in relazione ai Big Data è il caso del portale di lancio di Ashley Madison, che è già stato menzionato come Esempio nel capitolo 2. In questo caso, i dati personali sono diventati pubblici e sono stati utilizzati per ricattare i proprietari di i dati.

Le normative e le leggi sulla protezione dei dati aiutano a proteggere i consumatori dagli abusi. La base della legge generale sulla protezione dei dati nell'Unione europea e in Austria è il regolamento generale sulla protezione dei dati, entrato in vigore il 25 maggio 2018.



Excursus

Il Regolamento Generale di Protezione dei Dati - GDPR

Il regolamento generale sulla protezione dei dati, o GDPR in breve, è chiamato nella sua interezza "Regolamento (UE) 2016/679 del Parlamento europeo e del Consiglio, del 27 aprile 2016, relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati e che abroga la direttiva 95/46/CE".

Questo regolamento consente ai cittadini dell'UE di controllare meglio la raccolta e l'uso dei propri dati personali. Ciò dovrebbe rafforzare la fiducia dei consumatori nelle singole società. I diritti esistenti dei cittadini dell'UE sono consolidati nel GDPR e vengono anche stabiliti nuovi diritti. I diritti stabiliti nel GDPR includono:

- **accesso semplificato ai dati personali** - ciò include la fornitura di informazioni complete, chiare e comprensibili sul trattamento dei dati
- un nuovo **diritto alla trasferibilità dei dati**: i dati personali saranno trasferiti in modo semplificato
- un **diritto più chiaro alla cancellazione** ("diritto all'oblio"): i dati vengono cancellati se un cittadino non acconsente al trattamento dei propri dati e non vi sono motivi legittimi per conservarli
- un **diritto ad essere informato sui dati personali compromessi**: le aziende e le organizzazioni informano immediatamente le persone interessate in merito a gravi violazioni della protezione dei dati personali. Anche l'autorità di controllo responsabile della protezione dei dati deve essere informata

Per le aziende, il GDPR ha lo scopo di creare maggiori opportunità commerciali e promuovere l'innovazione con numerose misure. Queste possono includere:

- la creazione di **norme uniformi a livello dell'UE**, che porteranno a importanti risparmi
- la **nomina di un responsabile della protezione dei dati** all'interno delle autorità e delle società che si occupano di set di dati di grandi dimensioni
- la **designazione di un unico punto di contatto** nel proprio paese a cui le imprese devono rivolgersi
- la creazione di **norme UE per le società di paesi terzi** alle quali le società di paesi terzi devono aderire quando offrono beni o servizi o monitorano il comportamento delle persone
- la creazione di **regole che promuovano l'innovazione** e garantiscano che le norme sulla protezione dei dati siano prese in considerazione nelle prime fasi dello sviluppo di servizi o prodotti
- l'uso di **tecniche compatibili con la protezione dei dati come la pseudonimizzazione** (sostituzione di passaggi in un set di dati che consenta di identificare la persona associata) e la **crittografia** (i dati sono crittografati in modo che possano essere letti solo da persone autorizzate)
- eliminare gli **obblighi di segnalazione** per le società al fine di promuovere la libera circolazione dei dati personali all'interno dell'Unione europea
- svolgere **valutazioni d'impatto** quando il trattamento dei dati può costituire una minaccia per i diritti e le libertà delle persone

Il GDPR completo può essere consultato a <https://eur-lex.europa.eu/legal-content/IT/ALL/?uri=CELEX:32016R0679>.

Un'ulteriore sfida è connessa al fatto che i dipendenti esistenti nelle aziende non hanno sempre le competenze necessarie e non sono aperti alle possibilità che l'analisi di grandi quantità di dati offre all'azienda.

Il tempo e le risorse sono spesso sprecati perché i soggetti coinvolti non hanno idee chiare sullo scopo di un progetto Big Data o su quale infrastruttura sia necessaria per esso. Trovare e trattenere dipendenti competenti è di solito difficile per le aziende, poiché sono molto richieste.



Inoltre, la tecnologia Big Data è diversificata e confusa per i principianti. Hai mai sentito parlare di Spark, Hadoop MapReduce, Cassandra o Hbase? Queste sono tecnologie Big Data con diverse funzionalità e vantaggi.

Inoltre, le tecnologie si stanno evolvendo rapidamente, quindi le aziende spesso non riescono a tenere il passo con la loro adozione. Pertanto, per le aziende che stanno valutando di utilizzare un'analisi dei Big Data, è utile la consulenza di esperti.

Un altro punto è che i progetti di Big Data sono molto costosi. Questo vale sia per le aziende che scelgono un modello locale sia per quelle che preferiscono un modello cloud. La differenza è che con un modello on-premise, l'azienda utilizza il software per big data nel proprio data center ed è responsabile del suo funzionamento e manutenzione. In un modello cloud, d'altra parte, il software viene affittato solo dall'azienda e i dati rimangono al fornitore.

Definizione

Modello On-Premise

... si riferisce a una soluzione in cui l'azienda acquista o noleggia software Big Data e lo distribuisce nel proprio centro dati. La società deve prendersi cura dell'hardware stesso e si assume anche la responsabilità dell'uso del software e i dati vengono archiviati presso l'azienda.

Definizione

Modello Cloud

... si riferisce a una soluzione in cui un'azienda acquista il software Big Data come servizio; il fornitore si assume la responsabilità della manutenzione e del funzionamento. La società paga un prezzo di noleggio che include i costi di hardware, funzionamento e manutenzione. Con questa soluzione, i dati vengono archiviati presso il provider.

Se un'azienda decide di utilizzare una soluzione locale, questa deve investire in nuovo hardware e assumere nuovi dipendenti per far funzionare e mantenere il sistema. Nel caso di una soluzione cloud, l'azienda deve solo assumere dipendenti per gestire e mantenere il sistema e l'azienda deve pagare per i servizi cloud.

Dopotutto, la qualità dei dati è spesso scarsa e le aziende devono affrontare la sfida di armonizzare i dati provenienti da diverse fonti di diversa qualità. Ad esempio, un commerciante online analizza i dati provenienti da social media, registri di siti Web, call center e siti Web che hanno formati diversi.

Ma anche quando tutti i problemi citati sono stati risolti, spesso non è così semplice per le aziende ottenere informazioni utili da grandi quantità di dati. Se le informazioni sono collegate tra loro e vengono tratte conclusioni errate, ad esempio, ciò può essere pericoloso.

Per Esempio, una persona può essere considerata non degna di fiducia da una banca che esegue un'analisi dei Big Data perché vive nello stesso quartiere di molte persone ritenute a rischio e guida la stessa macchina di molte persone considerate a rischio. Il seguente esempio mostra anche perché l'uso corretto di grandi quantità di dati è cruciale:

Esempio

Un rivenditore online si affida alla Big Data Analytics, che si basa su dati storici sul comportamento dei clienti. Si scopre che le persone che acquistano scarpe da ginnastica nere spesso aggiungono un paio di calze da ginnastica nere. Il rivenditore regola di conseguenza la sua scelta di una nuova gamma per la primavera. Tuttavia, poco prima dell'inizio della primavera, un noto rapper pubblica una sua foto con scarpe da ginnastica nere e calze gialle su Instagram. Molti giovani sono quindi alla ricerca di calze gialle da abbinare alle loro sneaker nere, ma sfortunatamente il rivenditore online si esaurisce presto perché non era preparato per la corsa. Il rivenditore ha semplicemente utilizzato la strategia sbagliata per i Big Data, basandosi solo sui risultati storici e ignorando altre importanti fonti di dati come social media, negozi di concorrenti, ecc..

Ricorda

In sintesi, queste sono le principali sfide che le aziende devono affrontare quando utilizzano i Big Data:

- garantire **la sicurezza dei dati** – in accordo con il GDPR
- **le competenze professionali** dei dipendenti – **uso competente** di tecnologie in rapido progresso
- **alti costi dei progetti Big Data** (hardware e software o costi di noleggio, personale, manutenzione ecc.)
- **scarsa qualità dei dati**, standardizzazione dei dati in diversi formati e con diversa qualità
- **corretta interpretazione** dei dati

Come hai notato, i Big Data offrono enormi possibilità e opportunità che le aziende non si sono nemmeno avvicinate allo sfruttamento. Tuttavia, i grandi volumi di dati sono anche associati a sfide e rischi che non devono essere sottovalutati e sono inquietanti per molte persone. Il fattore decisivo per garantire che i Big Data vengano utilizzati con successo senza causare danni ad altre persone è quindi la gestione responsabile e competente dei grandi volumi di dati.



3.6 Sommario

I Big Data si riferiscono a grandi quantità di dati che non possono più essere elaborati con software o hardware convenzionali e per i quali l'elaborazione e l'analisi vengono eseguite per uno scopo specifico. A differenza dei Big Data, Small Data si riferisce a dati accessibili agli esseri umani a causa del loro volume e formato.

Incontriamo queste grandi quantità di dati in situazioni quotidiane, ad esempio durante la navigazione sui social media o la ricerca su Google. Per definire meglio i Big Data, l'analista Doug Lane ha progettato il modello a 3 V, che afferma che i Big Data sono caratterizzati da volume, velocità e diversità.

Grandi quantità di dati possono essere utilizzate, tra l'altro, per migliorare l'orientamento strategico delle aziende, per i sistemi di apprendimento profondo, per combattere la criminalità e il terrorismo, per le indagini scientifiche sui fenomeni naturali (ad esempio terremoti e cambiamenti climatici), ma anche per valutazioni illegali che possono portare a ricatti o frodi. Il fattore decisivo non sono tanto i grandi volumi di dati stessi, ma ciò che accade.

Le aziende possono utilizzare i Big Data per aumentare il successo aziendale. Tra l'altro, Big Data Analytics consente alle aziende di prendere decisioni aziendali migliori e valutare i rischi con maggiore precisione. Inoltre, l'efficienza dei processi aziendali può essere aumentata quando i dati vengono analizzati, valutati e collegati tra loro. I Big Data aiutano le aziende in ricerca e sviluppo a fare previsioni su tendenze, caratteristiche del prodotto, ecc. Infine, le conoscenze acquisite dai Big Data possono essere utilizzate anche per offrire un servizio clienti personalizzato.

Per un'analisi dei Big Data di successo, sono richiesti un'adeguata strategia di Big Data, una cultura aziendale adeguata, personale con il necessario know-how, tecnologia efficiente e, ultimo ma non meno importante, un'adeguata strategia di protezione dei dati. Ma analizzare e trattare i big data non solo offre opportunità e possibilità, ma pone anche sfide e rischi.

Una grande sfida per le aziende è garantire la sicurezza dei dati e conformarsi al regolamento generale sulla protezione dei dati. Inoltre, è spesso difficile per le aziende trovare e conservare professionisti idonei in grado di gestire la complessa tecnologia dei Big Data. I progetti di big data sono anche associati a costi elevati e la qualità dei dati è spesso scarsa. Infine, le giuste conclusioni devono essere tratte dai risultati dell'analisi dei dati e devono essere prese le giuste decisioni.