

# Great Data. By Design.



I Great Data non sono casuali.  
Si ottengono grazie alla progettazione.

Non è facile assicurarsi di disporre dei dati bonificati, sicuri e connessi di cui si ha bisogno per alimentare decisioni sicure e processi di business efficaci.

**Bisogna lavorarci...**

A photograph of a server room. In the center, a man wearing a light-colored shirt and dark trousers stands on a white floor, facing away from the camera. He is holding a laptop and appears to be working on it. Behind him is a large server rack filled with various electronic components and cables. The room has a high ceiling with exposed pipes and ductwork. The lighting is bright, creating a hazy effect around the man.

La difficoltà è che le tendenze del mercato stanno remando contro gli utenti e i professionisti dei dati.

I lavori stanno diventando  
più difficili.

Tendenza del mercato

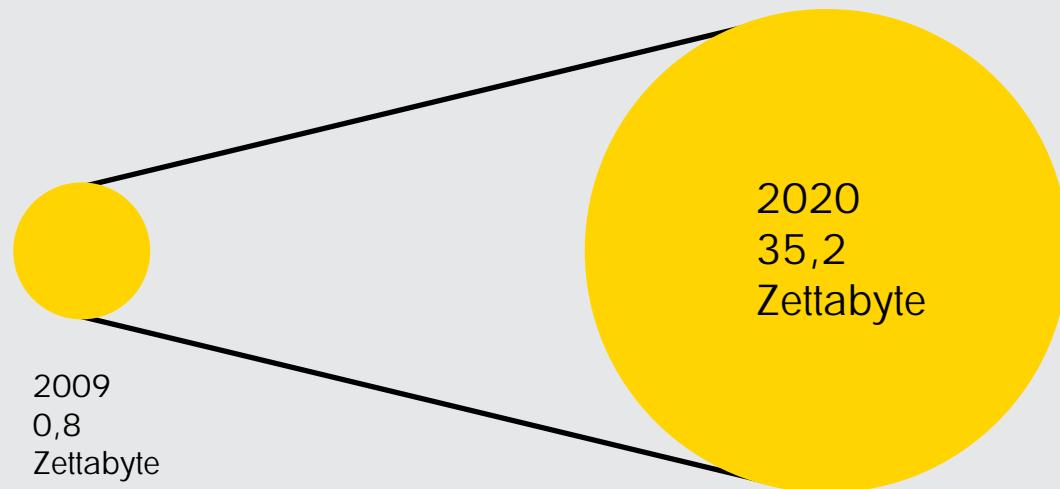


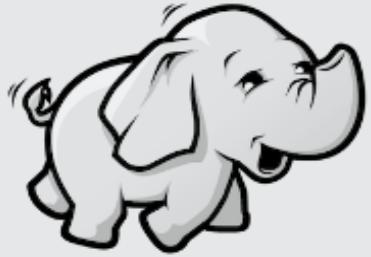
#1

Più dati. In più luoghi. Che si muovono più veloci che mai.

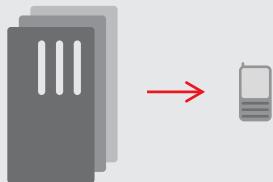
Il volume, la velocità e la varietà dei dati stanno aumentando con un ritmo senza precedenti. La quantità di dati generati nel mondo oggi raddoppia ogni due anni.

È la nuova legge di Moore.





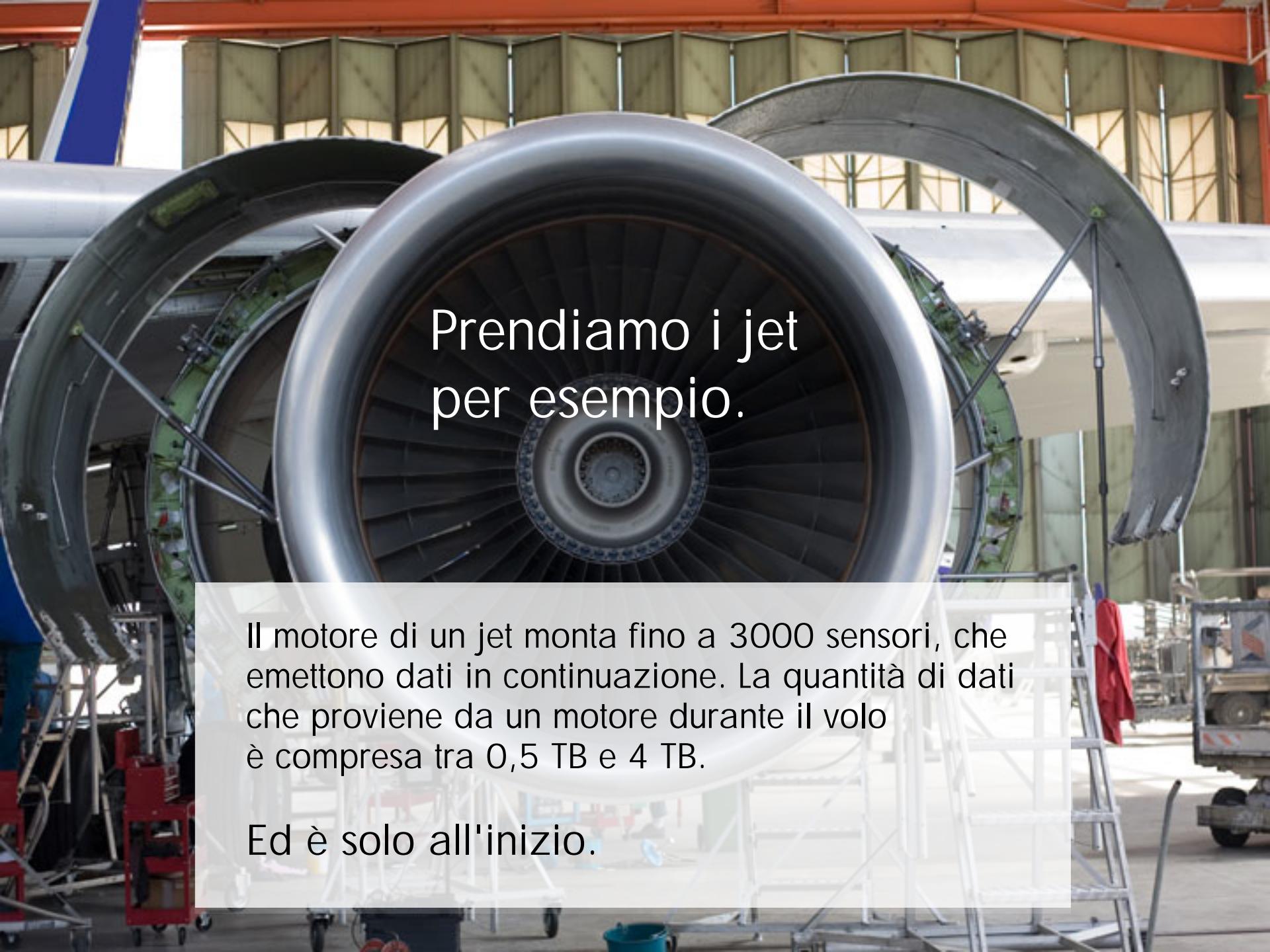
Le tecnologie come **Hadoop** ci offrono una soluzione accessibile per archiviare grandi quantità di dati.



La potenza del **mainframe computing** ora può stare **nel palmo di una mano**.



E, per completare l'opera, stiamo collegando dispositivi e sensori RFID a tutto quanto.



Prendiamo i jet  
per esempio.

Il motore di un jet monta fino a 3000 sensori, che emettono dati in continuazione. La quantità di dati che proviene da un motore durante il volo è compresa tra 0,5 TB e 4 TB.

Ed è solo all'inizio.

Il volume, la varietà e la  
velocità di dati non potranno  
che aumentare.

Tendenza del mercato

#2

I dati sono ovunque  
e la loro qualità è discutibile



I dati sono  
sparsi  
ovunque



File flat



Dispositivi mobili



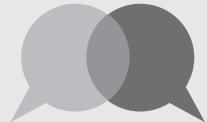
Database



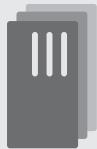
Sistemi legacy oscuri



Sensori



Social media



Mainframe



Applicazioni CRM



File flat



Cluster Hadoop



Cloud



Applicazioni ERP



Code messaggi



Documenti non strutturati

Sono in tutti i luoghi tradizionali e in tutti quelli nuovi.  
Sia on-premise sia nel Cloud.

A photograph showing a long row of server racks in a data center. The racks are dark grey or black and filled with various electronic components and cables. The floor is made of grey tiles, and the ceiling above is white with some structural elements visible.

In genere i progetti di data integration erano limitati o messi a rischio dal costo e dalle performance della CPU, della memoria, della rete o del disco. Oggi non è più così.

Ora siamo limitati dalla nostra capacità di trattare dati frammentati e di qualità scarsa o discutibile.

Per valorizzare a pieno i loro dati, le aziende devono essere in grado di integrarli in tutta l'impresa.



E la data quality deve essere integrata nel processo. Proprio come la produzione ha attraversato una transizione negli anni '80, quando i passaggi relativi alla qualità per la realizzazione di prodotti vennero integrati *nel* processo di produzione, è necessario fare lo stesso con i dati.

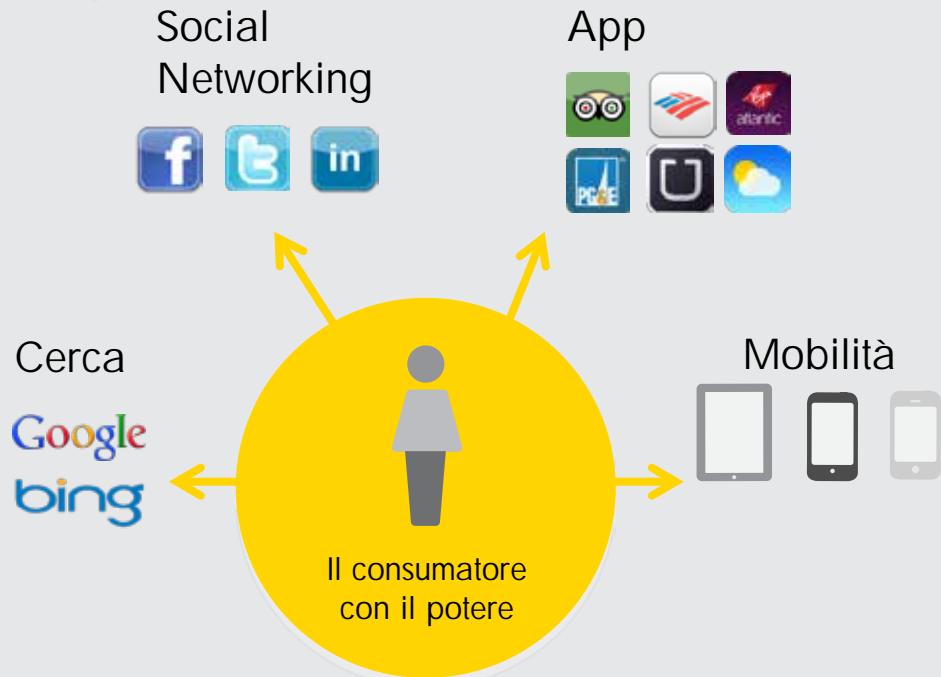
Tendenza del mercato



#3

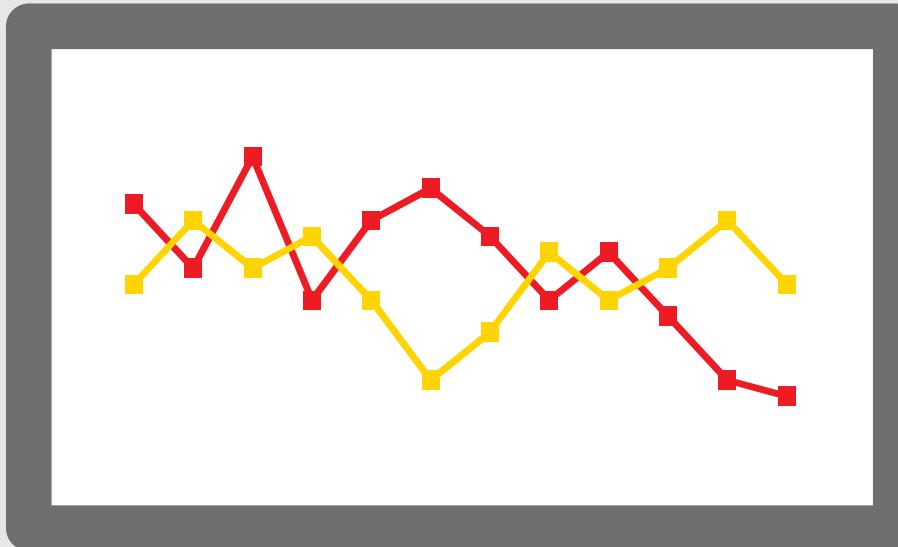
Il business chiede il self service

Nel corso degli ultimi cinque anni, gli utenti di business hanno aumentato le proprie competenze tecniche. La tecnologia di facile uso ora ha un ruolo importante nelle loro vite personali e li aiuta a fare le cose più velocemente, con maggiore facilità e in modo migliore. Ha dato loro maggior potere. Ed essi si aspettano la medesima esperienza al lavoro.



Il dipartimento commerciale non è disposto ad aspettare che il dipartimento IT consegni i Great Data. Vuole farlo per conto proprio.

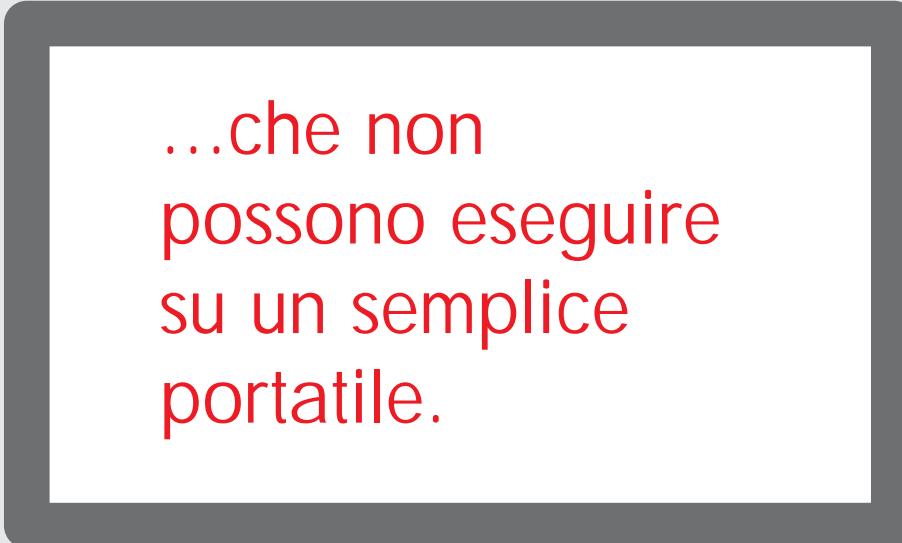
Ci sono alcuni ottimi tool di self-service che consentono loro di visualizzare i dati.



Il problema è che funzionano solo per un singolo set di dati alla volta.

Quando il dipartimento commerciale ha bisogno di dati che attraversano i confini aziendali, o quelli dei set di dati, deve comunque rivolgersi al dipartimento IT.

O ancora peggio, si rivolge a esso perché ha fatto tutto ciò che poteva con i tool self-service, per poi rendersi conto che i dati che sta utilizzando sono mission-critical e richiedono processi mission-critical...



...che non  
possono eseguire  
su un semplice  
portatile.

Il self-service può portare il  
dipartimento commerciale  
solo fin qui.



È necessario un  
nuovo modo di  
pensare

Molte aziende credono che il modo per ottenere un vantaggio competitivo sia concentrarsi sui processi del proprio core business.



Se siamo i migliori in  
quello che facciamo,  
possiamo battere la  
concorrenza.

E non hanno del  
tutto torto.

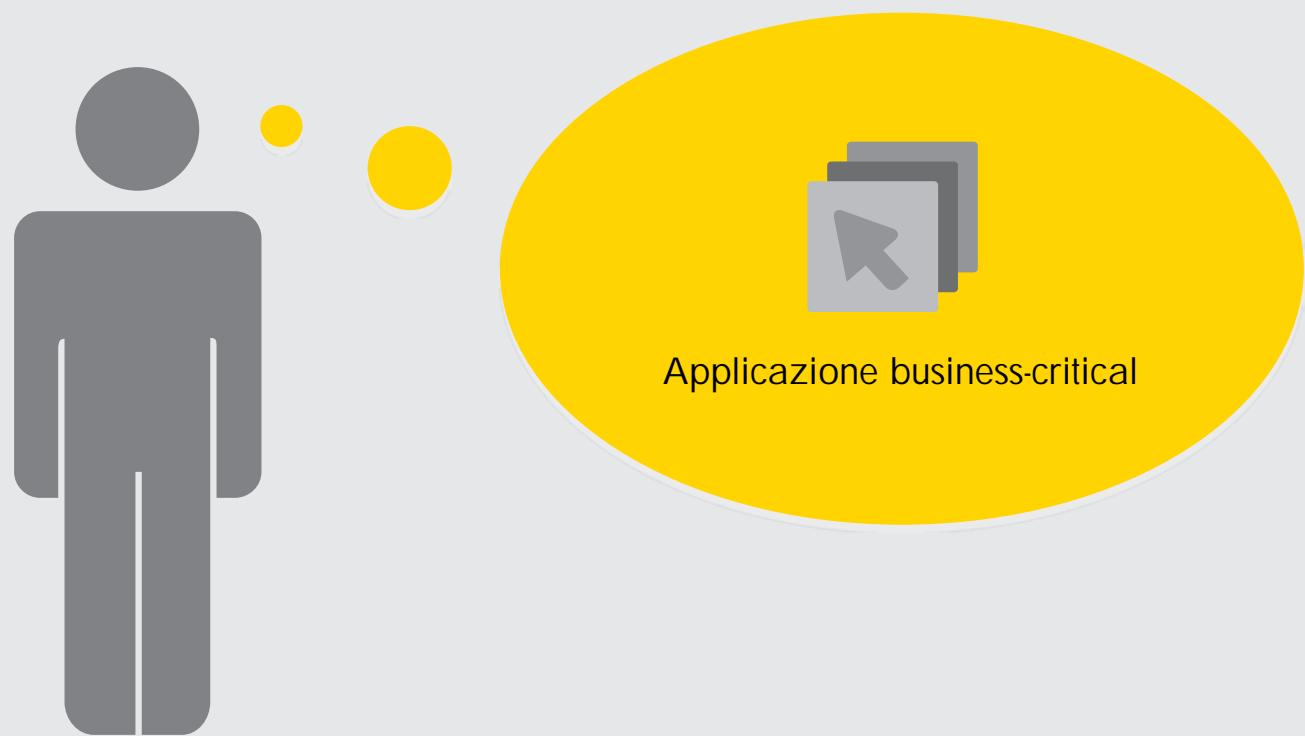
Sono convinti che investendo in applicazioni che supportano questi processi del core business potranno utilizzare le nuove efficienze, o il servizio migliorato che deriva da queste, per ottenere un vantaggio competitivo.



Abbiamo bisogno di un'applicazione che automatizzerà e migliorerà i nostri processi centrali, in modo da poter battere la concorrenza.

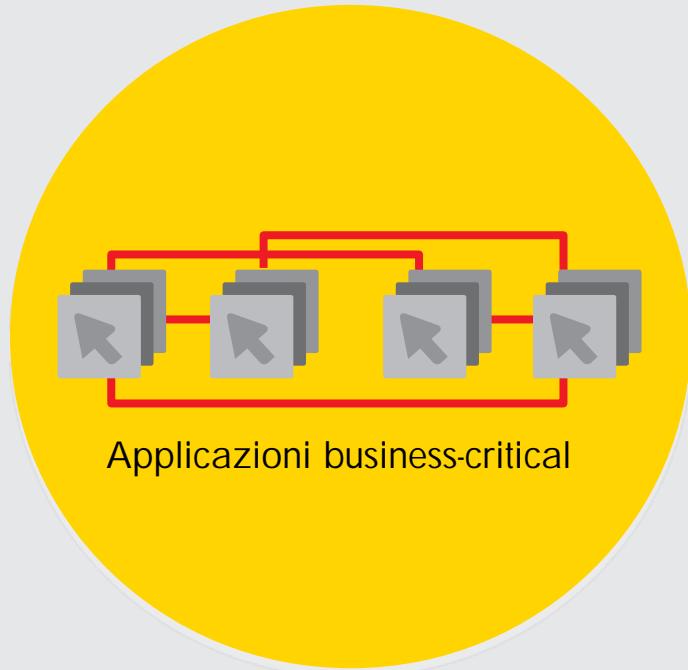
E non hanno del tutto torto.

Il problema è: le persone continuano a pensare alla propria applicazione di business come a un singolo elemento monolitico.

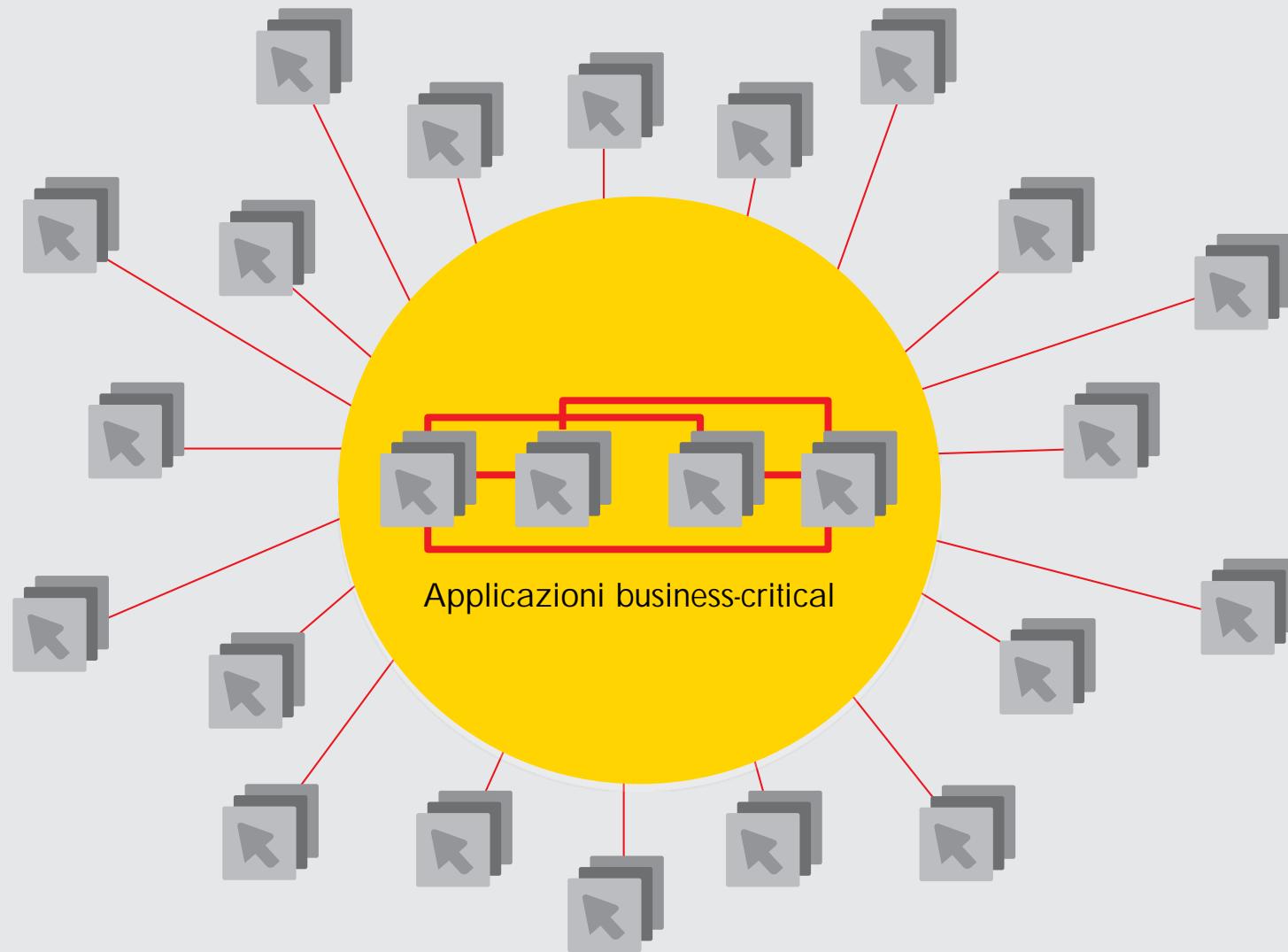


È qui che sbagliano.

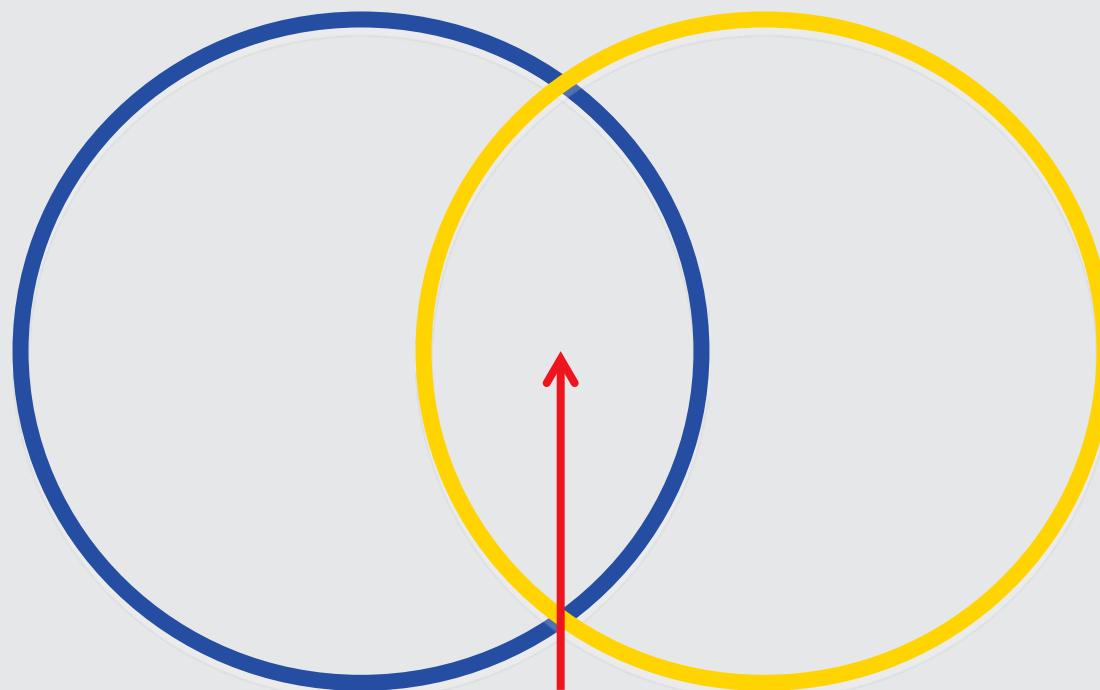
La realtà è che questi processi e le applicazioni fondamentali che le supportano non sono un singolo elemento monolitico. Qualunque processo di business oggi è fortemente distribuito su più sistemi...



.... e il numero di sistemi e di punti di dati verso o da cui i dati devono fluire è soltanto in aumento.



È vero in generale che l'innovazione esiste ai margini dei confini, o all'intersezione di diverse discipline.



L'innovazione avviene qui

Con la continua creazione di dati che attraversano più sistemi, la capacità di integrare e intersecare i dati attraverso questi confini diventa un fattore di successo critico per l'innovazione di nuova generazione.

Come si accelera la supply chain?

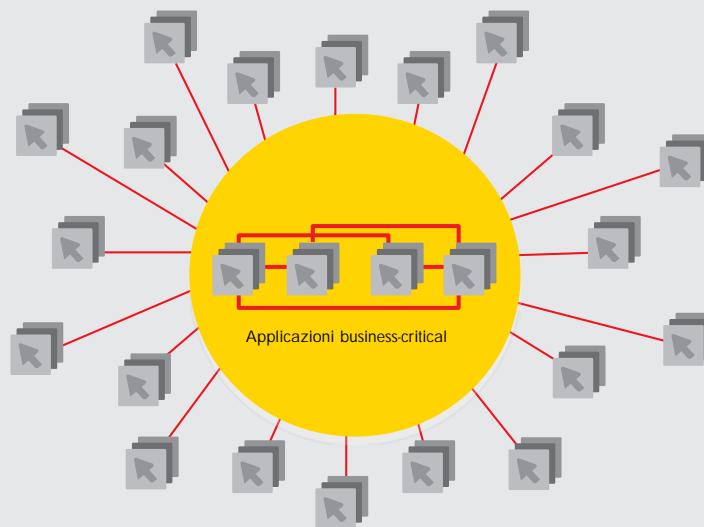
Possiamo favorire le efficienze nei nostri processi di approvvigionamento?"

Come si può migliorare la collaborazione tra vendor e appaltatori?"

Chi sono i nostri clienti più profittevoli?

Possiamo creare nuovi servizi basati sulle informazioni da offrire ai nostri clienti?

Abbiamo tutti i dati di cui abbiamo bisogno per supportare i nostri vincoli di conformità?



Ma integrare i dati è più  
difficile di quanto si possa  
pensare comunemente.



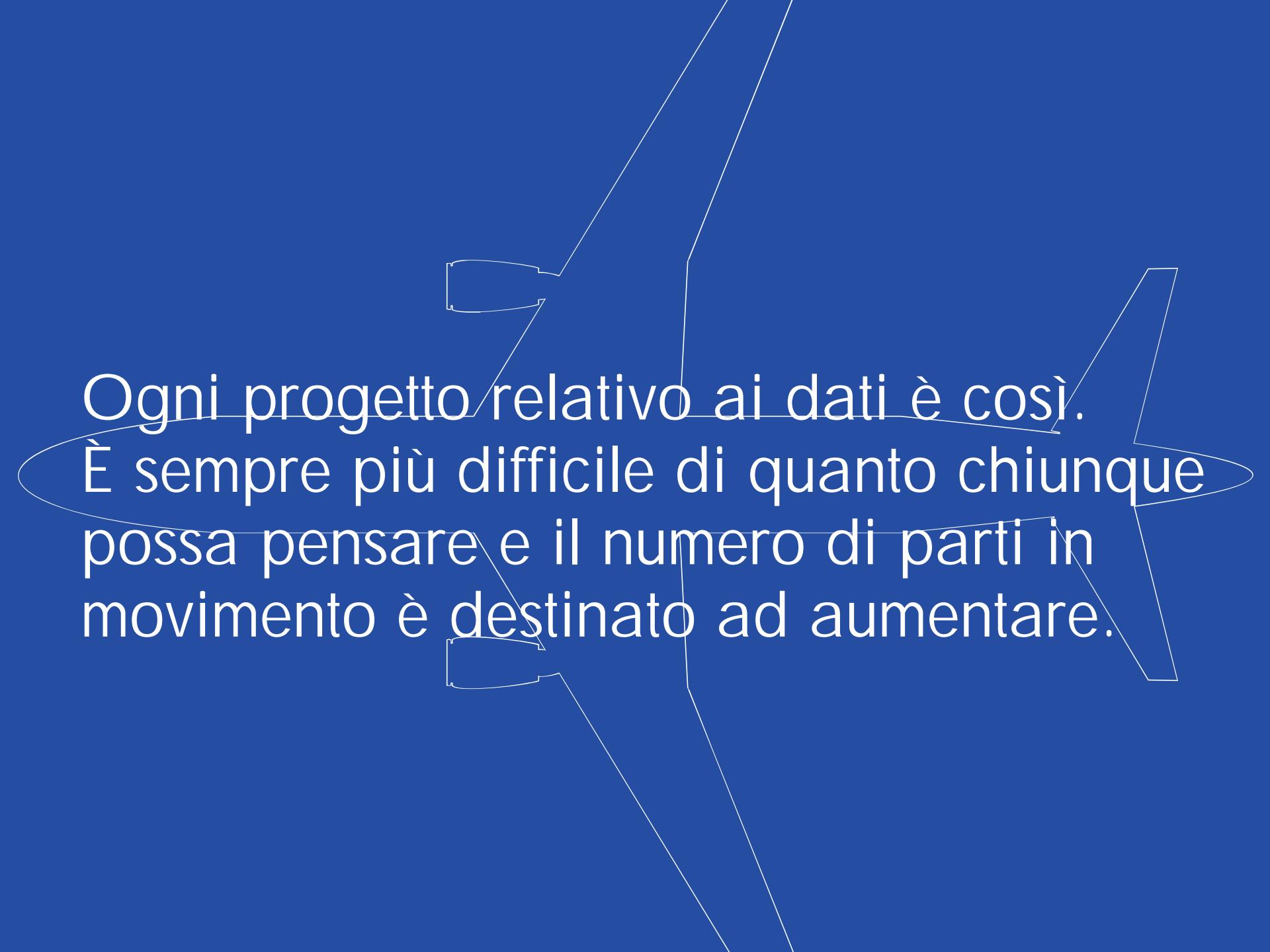
Prendiamo il jet, per esempio. Benché i motori possano essere gli stessi da un aereo all'altro, i dati provenienti da essi, attraverso i 3000 sensori, non sono controllati dal produttore del motore. Sono controllati dalle compagnie aeree.

E ciascuna compagnia aerea archivia questi stessi 3000 attributi nel proprio formato.

A white commercial airplane is shown from a low angle, parked at a grey concrete gate. The tail section and part of the fuselage are visible. The background consists of vertical grey panels of the airport terminal.

Questo significa che, quando i dati per lo stesso tipo di motore vengono inviati al produttore per l'analisi, questo per prima cosa deve normalizzarli. Quello che sembrerebbe un esercizio facile, analizzare i dati provenienti dallo stesso tipo di motore, è molto più difficile di quanto sembri.

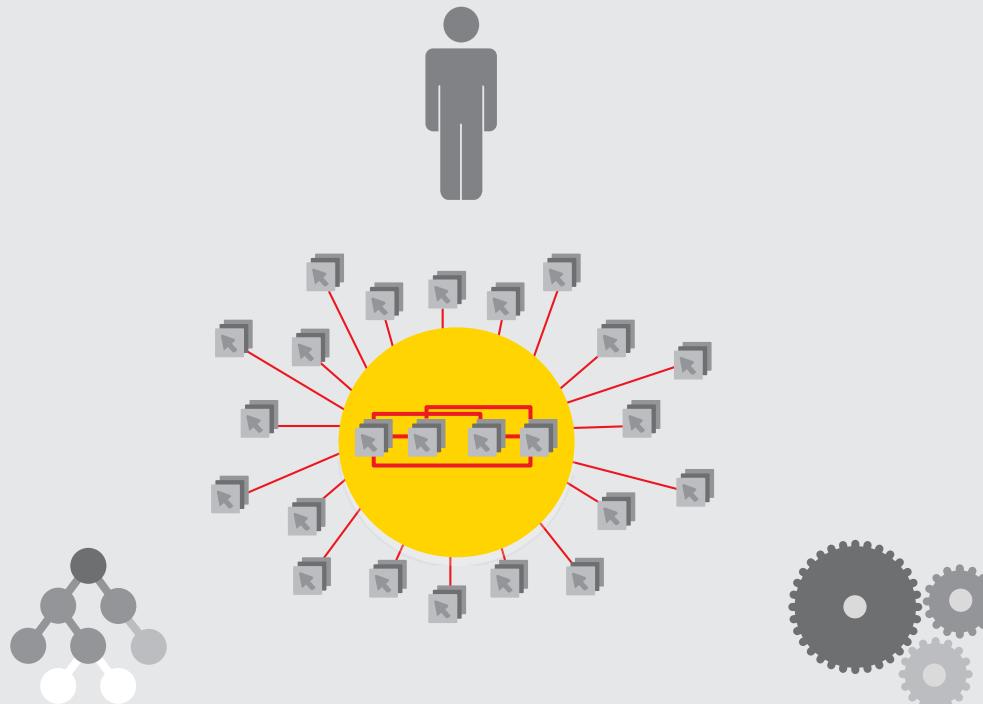
La difficoltà aggiuntiva è che i dati legacy non muoiono mai e devono essere inseriti anch'essi.



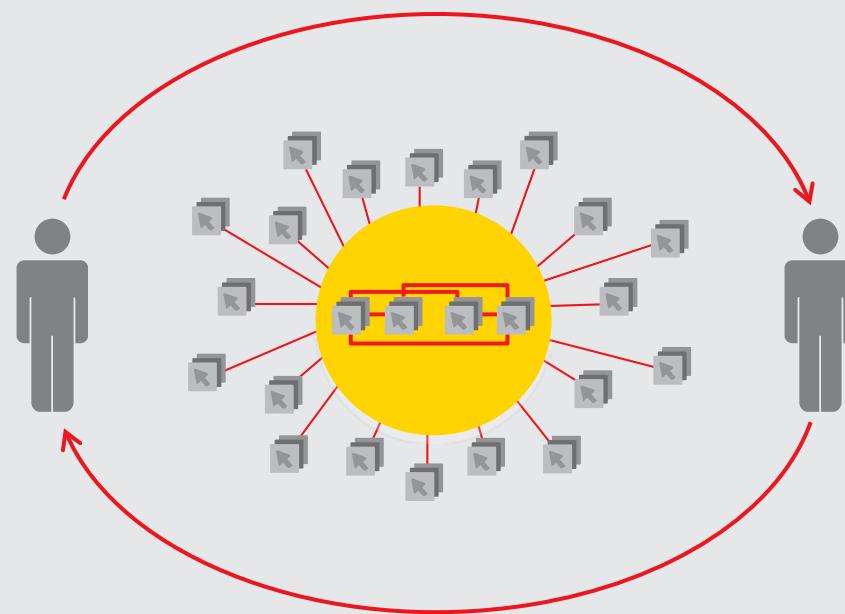
Ogni progetto relativo ai dati è così.  
È sempre più difficile di quanto chiunque  
possa pensare e il numero di parti in  
movimento è destinato ad aumentare.

Per superare questa  
difficoltà, è necessario  
progettare dei Great Data  
nei propri processi di  
business.

Proprio come si investe in persone, processi e tecnologia per i propri processi di core business, è necessario investire in persone, processi e tecnologia per integrare i dati distribuiti che supportano questi processi.



Questo perché l'agilità del business ora dipende dall'agilità della data integration. E quest'ultima dipende dal riuscire a coinvolgere tutti, e assicurarsi che dipartimento commerciale e IT dispongano dei tool adatti per consentire la collaborazione. Infatti, abbiamo visto che nelle aziende in cui dipartimento commerciale e IT collaborano, i progetti di DI vengono eseguiti a una velocità cinque volte superiore rispetto a quelle in cui ciò non avviene.



# 5

## Considerazioni per la designazione di Great Data



#1

## Collegare tutti i propri dati

RDBMS, file flat, XML, Hadoop, NOSQL, social media, mainframe, dati macchina e altro ancora ...

La data integration consente di combinare i dati provenienti da molte fonti diverse di rich data per produrre nuove informazioni di business che non si potrebbero ottenere da una singola fonte. Occorre assicurarsi che i propri tool di data integration siano in grado di connettersi con qualunque fonte dati (sia corrente sia legacy), compresi RDBMS, NOSQL, mainframe, testo, applicazioni e così via, e non solo alle fonti dati utilizzate attualmente. È questo insieme universale di connessioni che rende possibile riunire tutti i dati.



#2

## Supporto del formato giusto e della latenza

in modalità batch, tempo reale, quasi tempo reale.  
Strutturati, non strutturati, semistrutturati.

Allo stesso modo in cui la data integration trae dati da molte fonti diverse, deve anche essere in grado di utilizzare vari e molteplici tipi di dati, comprese fonti dati strutturate, semistrutturate e non strutturate in modalità batch e in tempo reale. È necessario un tool che sia sufficientemente flessibile per lavorare con qualunque tipo di dato in cui ci si imbatta.



#3

## Comprendere la struttura dei dati e del contenuto

Includere la profilazione dei dati nella metodologia

Con un numero così elevato di fonti dati diverse coinvolte, è necessario avere un mezzo per assicurarsi che i dati siano ciò che ci si aspetta. È importante che i tool consentano un livello di profilazione dei dati, in modo che sia possibile verificare i dati che entrano ed escono dal proprio sistema, e che si finisca con i risultati desiderati.



#4

## Favorire una collaborazione efficace tra dipartimento commerciale e IT

### Essere agili e lean

Non ci si può più più permettere di creare ed eseguire progetti utilizzando i tradizionali metodi di sviluppo isolati. I tool di data integration devono supportare processi di integrazione lean e agili che permettano la collaborazione tra dipartimento commerciale e IT, in modo che lo sviluppo avvenga in modo rapido e interattivo.



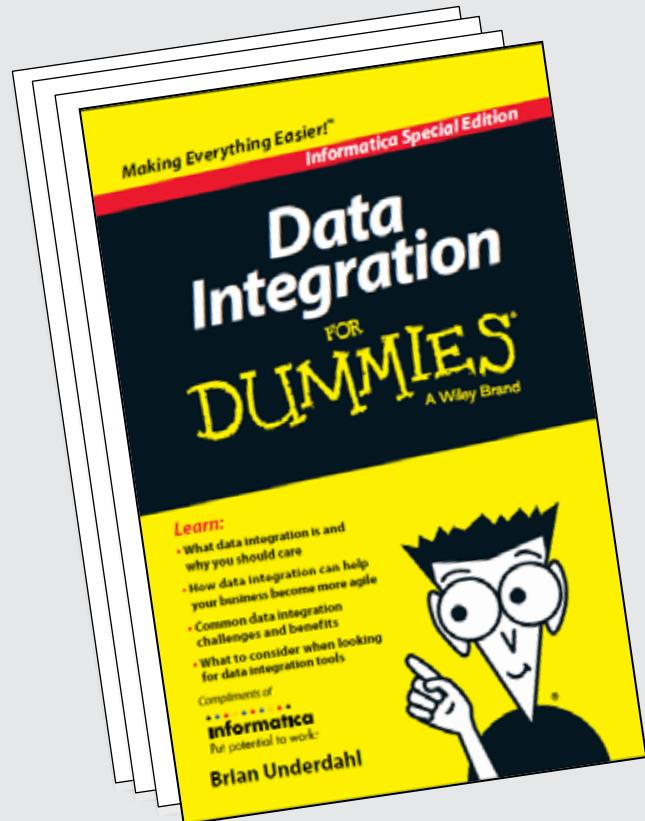
#5

## Supportare la crescita e l'espansione del business

Capacità di scalare verso l'alto e verso il basso

Le aziende crescono, e così anche le dimensioni dei loro progetti. Non si vuole certo rimanere bloccati in tool che sono appropriati solo per i progetti odierni. Si vuole invece avere la capacità di scalare, crescere e spostare i progetti dai piccoli esercizi di innovazione di singoli dipartimenti ai grandi ambienti mission-critical di livello enterprise, o viceversa.

# Scoprire come **si** possano costruire Great Data, come intento progettuale.



[Fai clic qui per scaricare](#)