

[Home](#) > [AI/ML](#)

L'importanza della Data Culture nell'era dei Business Data-Driven

7 Giugno 2024 - 7 min. read

[AI](#) [Data and Analytics](#) [Data Ingestion](#) [Data Security and Governance](#)

[Data Visualization](#)

Nel mondo digitale di oggi, i dati svolgono un ruolo fondamentale.

Consentono alle aziende di restare competitive rispondendo velocemente all'evoluzione dei propri mercati di riferimento grazie a decisioni informate e migliori strategie di business sempre più orientate al miglioramento dell'esperienza utente.

L'utilizzo del dato, tuttavia, può fare la differenza solo se calato all'interno di una più ampia e condivisa Cultura aziendale del Dato.

Instillare efficacemente una Data Culture in azienda è un processo lento e complesso che deve necessariamente coinvolgere tutti gli stakeholder dell'organizzazione, dal management, ai decision maker, arrivando poi a tecnici come i data analyst.

Ma cosa si intende per Cultura del Dato nell'era del data-driven? E perché è fondamentale averla per passare con successo da un approccio più tradizionale a uno basato sui dati?

Ne parliamo in questo articolo soffermandoci sugli aspetti imprescindibili su cui ciascuna azienda dovrebbe concentrarsi.

Informazioni VS Dati

Partiamo da una definizione che ci aiuta ad avere fin da subito una visione corretta sul tema.

Affinché si possano prendere decisioni migliori, serve ottenere informazioni, ovvero qualcosa a cui siamo in grado di attribuire un significato. Le informazioni non nascono tali, ma devono essere estrapolate elaborando e interpretando i dati, cioè i tasselli che le compongono.

“Dato” e “informazione” quindi non sono esattamente sinonimi: non tutti i dati, come vedremo più avanti nell’articolo, potranno diventare informazioni, ma avere dati è condizione necessaria per poter dedurre informazioni.

Data Sourcing

Fatta questa premessa, è chiaro che la prima cosa che un’azienda deve fare è mettersi in condizione di raccogliere dati. Ovviamente è impensabile - e poco sensato - raccogliere tutto fin da subito.

Il nostro consiglio è quindi di partire con una **mappatura dettagliata** di tutto ciò che può generare dati (applicazioni, macchinari, ma anche sonde, sensori e interazioni utente in generale) e strutturare poi la raccolta di conseguenza.

Quanti dati servono per un progetto data-driven di successo?

Nel mondo Data è praticamente impossibile immaginare fin dal giorno zero tutti i possibili scenari di un progetto; mai come in questo caso, l’appetito vien mangiando!

In quest’ottica, l’abbondanza di dati è spesso un vantaggio.

I dati però possono acquisire valore solo se vengono storicizzati, messi in relazione tra di loro e archiviati all’interno di un unico Data Lake contenente un set di dati sufficientemente grande.

Da questo punto di vista, le possibilità offerte dal Cloud sono moltissime: in particolare, logiche oculte di storage tiering permettono di archiviare ad un costo contenuto grandissime quantità di dati non solo per prendere decisioni sul breve periodo, ma anche per un vantaggio competitivo futuro e un go-to-market più veloce. Anche senza avere fin da subito un’idea precisa di quanti e quali dati ci serviranno per raggiungere un determinato obiettivo di business.

Possedere tanti dati - per di più senza preoccuparsi del costo -, tuttavia, non basta per assicurarsi un buon risultato. Ci sono infinite variabili che concorrono, soprattutto quando - come dovrebbe sempre essere - si ha l'idea di sfruttare i dati per fare qualcosa di pionieristico.

È frequente, ad esempio, accorgersi in corso d'opera che la domanda iniziale a cui si voleva trovare risposta attraverso l'uso dei dati non è quella giusta in relazione ai dati che si hanno effettivamente a disposizione o, ancora, che la risposta ottenuta dalle informazioni raccolte non è ciò che ci saremmo aspettati.

Per esplorare i dati e comprenderne le potenzialità, è consigliato effettuare un'**analisi esplorativa dei dati (EDA)**, che consente di visualizzare le loro caratteristiche principali attraverso l'uso di grafici e statistiche descrittive. Questo aiuta a individuare **pattern**, **anomalie** e **relazioni tra variabili**.

Strumenti come Python e R, con librerie quali Pandas, NumPy, Matplotlib, e Seaborn, sono essenziali per eseguire EDA in modo efficiente e approfondito. Inoltre, strumenti che aiutano i Data Analyst come i notebook Jupiter possono rendere l'analisi rapida ed efficace.

Un'altra tecnica fondamentale è il data mining, che sfrutta algoritmi avanzati per scoprire pattern nascosti nei dati. Strumenti come Apache Spark, Hadoop e SAS sono utilizzati per processare grandi volumi di dati e applicare algoritmi di clustering, classificazione e associazione.

Il nostro consiglio è quello mantenere sempre **un feedback loop breve** tra quando si inizia la raccolta del dato e quando si analizza il risultato della loro elaborazione e chiedersi spesso:

“cosa raccontano i dati che raccogliamo?”

In questo modo, si possono intercettare indicatori di business forti a cui non avevamo pensato e che potrebbero portare anche ad aggiustamenti vantaggiosi nella strategia commerciale.

Questo approccio iterativo e incrementale va di pari passo con il principio del fail-fast tipico del Cloud. Il basso costo del fallimento è una leva innovativa importantissima in quanto permette alle aziende di sperimentare in modo sostenibile percorrendo più

strade, anche in parallelo, ottenendo risultati in tempi brevi e contenendo enormemente anche i rischi.

Attenzione però: l'eccesso di informazione, distrugge l'informazione stessa! Tra le infinite possibilità di un progetto data-driven, quindi, come identificare un buon obiettivo di business e strutturare di conseguenza la raccolta di dati interessanti per centrarlo?

Questa è senza dubbio la parte con più incognite. Purtroppo, visione strategica aziendale e un buon team di data analyst ancora non bastano. Il supporto di un partner tecnologico fidato con profonda esperienza in ambito Cloud e progetti data end-to-end sapranno aiutarvi a costruire una buona data strategy.

Data Governance, Security e Compliance

Pur sembrando un passaggio semplice e in qualche modo “ovvio”, uno dei problemi più grandi per le organizzazioni è creare una Data Platform per rendere fruibili i dati a più professionalità in modo sicuro ed efficiente.

Una volta che abbiamo raccolto i dati e ne abbiamo verificato l'utilità in funzione dell'obiettivo, è il momento di utilizzarli per estrarne del valore ed è a questo punto che ci si inizia a scontrare con i primi problemi di **governance, compliance e sicurezza**.

Bisogna iniziare a gestire correttamente i permessi per assicurare l'informazione giusta alla persona giusta, al momento giusto. Bisogna definire le regole di accesso, identificare le entità che potranno operare sui dati e le regole tramite cui ciascuna di esse potrà farlo, e catalogare tutte le informazioni a seconda che siano dati pubblici o dati sensibili.

Anche in questo caso, esistono varie strategie. Quella che generalmente consigliamo per i casi in cui la situazione non sia troppo complessa è catalogare tutti i dati come **dati sensibili**, magari già mascherati e anonimizzati. In questo modo, avremo pochi soggetti autorizzati ad accedervi e quindi un rischio notevolmente ridimensionato in termini di sicurezza, controllo e utilizzo incompatibile con le policy aziendali.

Nelle organizzazioni molto ampie, invece, in cui coesistono molti stakeholder che possono trarre vantaggio dall'utilizzo dei dati, considerare tutto come “dato sensibile” può essere limitante. Una delle migliori strategie in questo caso è effettuare il **labelling**

dei dati in fase di raccolta classificandoli già correttamente in base al grado di riservatezza del dato.

L'uso di meccanismi di autenticazione multi-fattore e di politiche di controllo degli accessi basate sui ruoli (RBAC), che definiscono chiaramente i permessi di ciascun utente in base alle loro responsabilità, è un'altra delle best practices consigliate per gestire correttamente e in sicurezza una grande quantità di dati.

Un ulteriore aspetto cruciale è la garanzia della qualità dei dati attraverso la definizione di politiche di verifica e pulizia dei dati (data cleaning). Ciò implica la creazione di procedure standardizzate per l'integrazione, la trasformazione e la validazione dei dati, assicurando che le informazioni all'interno del data lake siano accurate, complete e affidabili.

Per garantire il mantenimento della qualità e della sicurezza è utile adottare strumenti di monitoring che possono aiutare a rilevare attività sospette e a mitigare i rischi associati.

Naturalmente, una volta che abbiamo messo i dati categorizzati da parte è necessario avere anche un modo efficiente per interrogarli, affinché le persone giuste possano effettivamente fruire dei dati di loro interesse.

Per Concludere

In questo articolo abbiamo approfondito l'importanza cruciale di **valorizzare i dati, renderli accessibili e integrarli all'interno del processo decisionale aziendale**, riconoscendo che queste azioni sono fondamentali affinché le informazioni possano trasformarsi in una reale risorsa competitiva. Abbiamo evidenziato come la consapevolezza del valore intrinseco dei dati sia il primo passo per trattarli in modo adeguato, garantendo che un giorno se ne possa estrarre valore.

La promozione di una cultura del dato richiede non solo la comprensione e l'adozione di nuove tecnologie, ma anche una **trasformazione del modo in cui l'intera organizzazione percepisce e utilizza i dati**. L'intera organizzazione deve essere coinvolta in questo sforzo collettivo, dal top management ai singoli dipendenti, al fine di superare le barriere esistenti e integrare pienamente il valore dei dati nella prassi quotidiana.

In conclusione, **l'adozione di una Data Culture richiede un approccio olistico, che combina la governance dei dati, la formazione e l'educazione continua del personale, e l'implementazione di strategie e strumenti adeguati.** Solo attraverso un impegno strategico e collettivo, le aziende possono trasformare i dati in un vantaggio competitivo sostenibile, affrontando con successo le sfide del cambiamento culturale e tecnologico.

About Proud2beCloud

Proud2beCloud è il blog di **beSharp**, APN Premier Consulting Partner italiano esperto nella progettazione, implementazione e gestione di infrastrutture Cloud complesse e servizi AWS avanzati. Prima di essere scrittori, siamo Solutions Architect che, dal 2007, lavorano quotidianamente con i servizi AWS. Siamo innovatori alla costante ricerca della soluzione più all'avanguardia per noi e per i nostri clienti. Su Proud2beCloud condividiamo regolarmente i nostri migliori spunti con chi come noi, per lavoro o per passione, lavora con il Cloud di AWS. Partecipa alla discussione!



Alessio Gandini

Cloud-native Development Line Manager @ beSharp, DevOps Engineer e AWS expert. Computer geek da quando avevo 6 anni, appassionato di informatica ed elettronica a tutto tondo. Ultimamente sto esplorando l'esperienza utente vocale e il mondo dell'IoT. Appassionato di cinema e grande consumatore di serie TV, videogiatore della domenica.
