

La digitalizzazione utile

Edoardo Vigo – Digitalization Manager
Sis.Ter – Centro di Trasferimento Tecnologico 4.0



Decidere con i dati

Big Data & Analytics

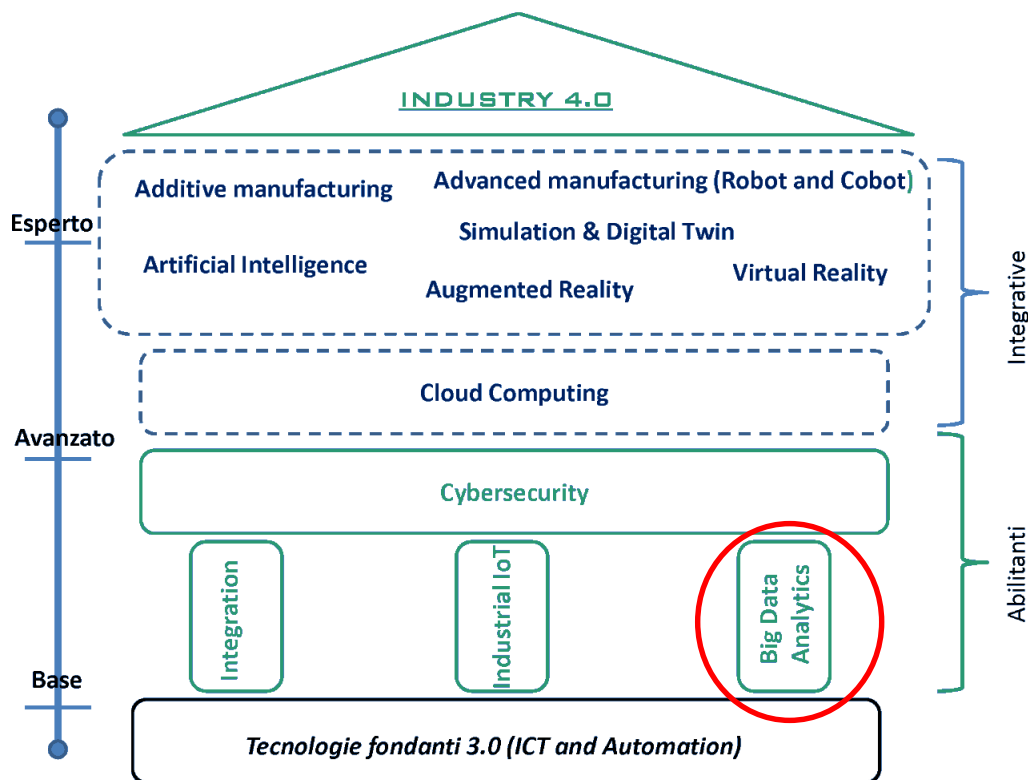
Archivio 4.0

16/03/2022

Agenda

- L'era del "dataismo"
- Che cosa sono i Big Data?
- Caratteristiche, tipologie e strutture
- Metodologie e tecniche (analytics)
- Esempi di applicazioni (Archivio 4.0)
- Opportunità derivanti dalla tecnologia
- Sfide della tecnologia e approcci da evitare
- Rubrica sui finanziamenti di progetti 4.0 (Livello Regionale)

Ancora i pilasti "abilitanti"



Diventa quindi determinate sfruttare tecniche e tecnologie che permettano alle imprese di creare valore da questa mole di dati.

Ed ecco il secondo pilastro abilitante del 4.0

Questa è decisamente l'era del "**dataismo**": le persone e le cose inviano continuamente dati a chi ha fornito il prodotto o il servizio.

I dati rappresentano oggi il maggior **agente trasformativo della società e delle imprese**.

Permettono di analizzare e comprendere una azienda come un **ecosistema che lega mercato a produzione**, un sistema di relazioni complesso che dovrebbe adattarsi alle dinamiche per **ottimizzare performance, creare valore**, rispondere ad **esigenze mutevoli**.

Small Data e Big Data

Gli *Small data* sono **dati sufficientemente “piccoli” per la comprensione umana**. Sono dati in un volume e in un formato che li rendono accessibili e fruibili rapidamente.

Sono quindi **piccoli set di dati che possono essere archiviati e processati da un normale pc**.



I *Big data* sono **“grandi” insiemi di dati caratterizzati da una crescita esponenziale, difficili da trattare con le soluzioni di archiviazione e analisi tradizionali**. Solitamente, come unità di misura si parla di **Petabytes (milioni di GBs)**.

L'evoluzione tecnologica e l'utilizzo che facciamo quotidianamente dei **vari devices (es. smartphone)** contribuisce in maniera evidente alla produzione di questa mole di **dati così elevata**.

Questo implica che più continuerà a crescere la quantità di dati a disposizione delle aziende, più **sarà necessario attuare processi tecnologici articolati per poterne poi estrarre informazioni utilizzabili**.

Smart Data

Gli Smart data sono, letteralmente, i **dati “intelligenti”**. Ovvero, i **dati selezionati, di elevata pertinenza, che contengono le informazioni utili a prendere decisioni.**

Gli Smart data sono la parte di Big data utile all'azienda, perché derivano dal processo di **selezione ed elaborazione “su misura”** rispetto al raggiungimento degli obiettivi precedentemente stabiliti.



I Big Data possono rendere **difficile aggregare, visualizzare e analizzare le informazioni in modo efficiente.**

Ecco perché è molto importante **contrastare questo diluvio di dati recuperando i dati necessari** che non solo **consentano di risparmiare risorse, ma che forniscano informazioni utili e preziose per l'azienda.**

Caratteristiche

I Big Data si possono quindi contraddistinguere per 4 caratteristiche:

Volume:
I dati hanno volumi
di crescita
esponenziali

Velocità:
I dati possono essere
aggiornati molto
rapidamente



Varietà:
I dati possono essere
strutturati, semi
strutturati e non
strutturati

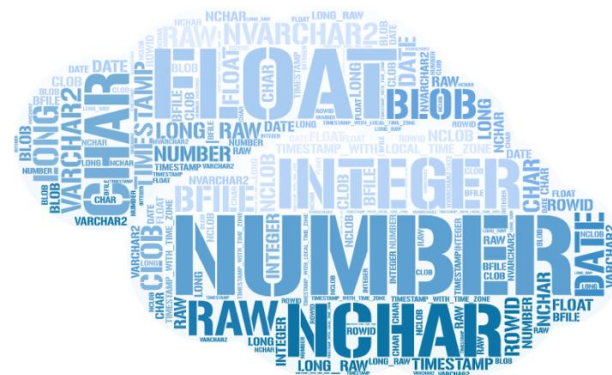
Veridicità:
I dati possono essere
più o meno affidabili
e attendibili

Una base dati allora sarà tanto più complessa quanto più sarà grande, più gli aggiornamenti saranno veloci e più i dati saranno poco strutturati e poco affidabili

Tipologie di dati

I dati possono presentarsi sotto diverse forme e tipologie (di seguito alcuni esempi):

- Dati alfa numerici
- Immagini/Video
- Audio
- Testi
- Serie temporali
- Dati Geo-spaziali
- Dati dei Social Media
- Dati IoT

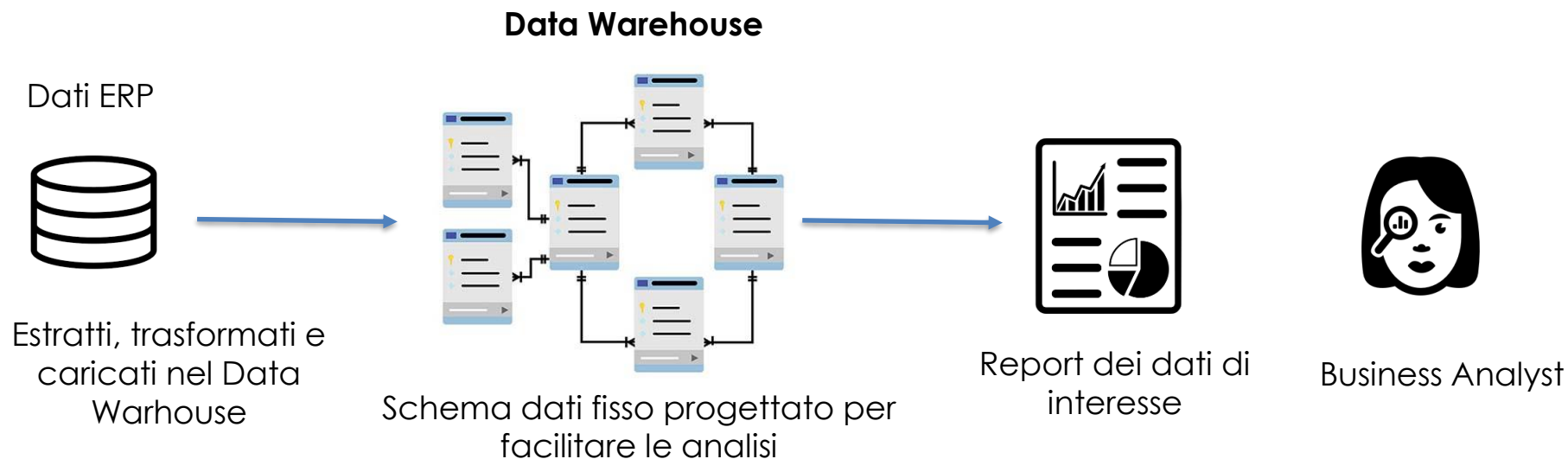


I Metadati

Un metadato letteralmente "(dato) per mezzo di un (altro) dato", è un **informazione che descrive un insieme di dati**.

Un esempio tipico di metadati è costituito dalla **scheda del catalogo di una biblioteca**, la quale contiene informazioni circa il contenuto, l'autore e la posizione di un libro, cioè dati riguardanti più dati che si riferiscono al libro. **Un metadato può essere anche un dato aggiunto all'insieme delle informazioni per altri scopi**. Ad esempio, se alla scheda del libro della biblioteca aggiungo un ID, ossia un identificatore univoco, quest'ultimo è un metadato.

Strutture (Data Warehouse)

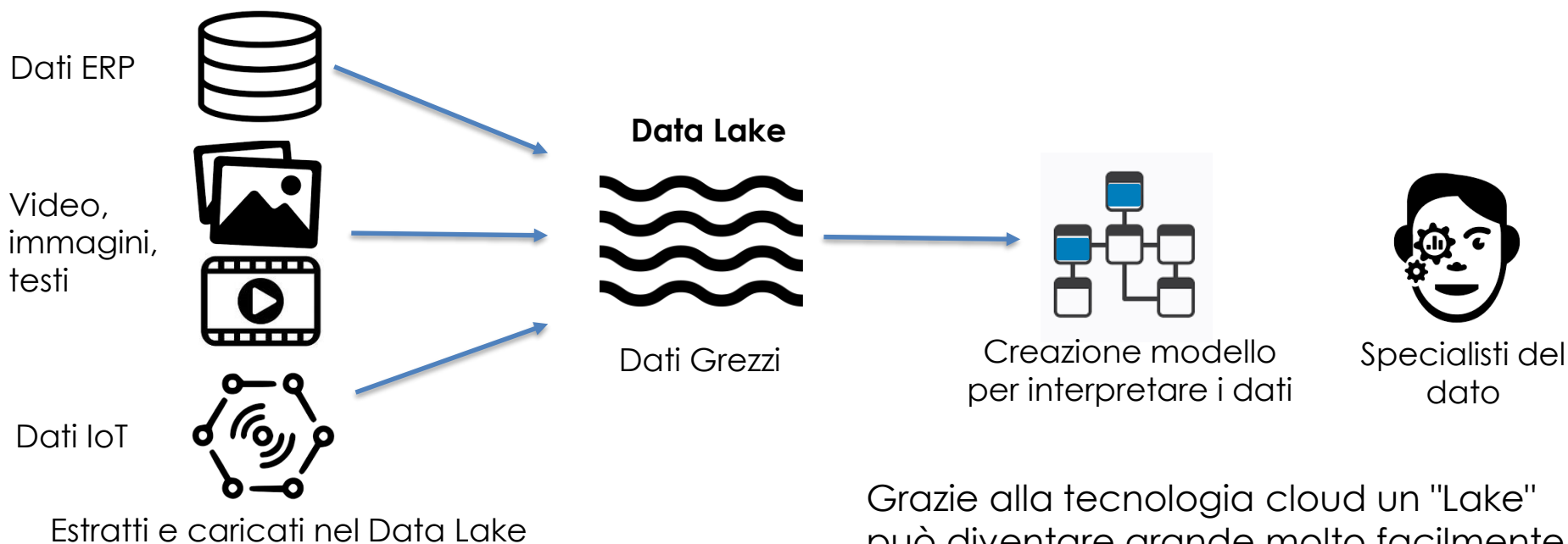


L'elaborazione analitica all'interno di un data warehouse viene eseguita su dati che sono stati preparati per l'analisi: raccolti, contestualizzati e trasformati, con lo scopo di generare insights basati sulle analisi. Sono specificamente progettati per **facilitare le analisi dati successive.**

Il Data Warehouse è la soluzione tradizionale per l'archiviazione di dati strutturati e semi-strutturati.

I SW di Business Intelligence e i Business analyst riescono a ricavare insights per l'azienda, oltre a mantenere un record storico.

Strutture (Data Lake)



I Data Lake **non hanno strutture dati fisse e preconfigurate** e possono ospitare qualsiasi tipologia di dato nel formato originale **senza doverli pre-elaborare**.

Soluzione **vantaggiosa per archiviare dati non strutturati**.

Grazie alla tecnologia cloud un "Lake" può diventare grande molto facilmente a seconda delle nostre esigenze (**gestisce bene i Big Data**).

Tipicamente è il contesto in cui i **Data Scientists si esprimono al meglio** potendo sperimentare per esempio algoritmi di ML.

Strutture e tecnologie che collaborano

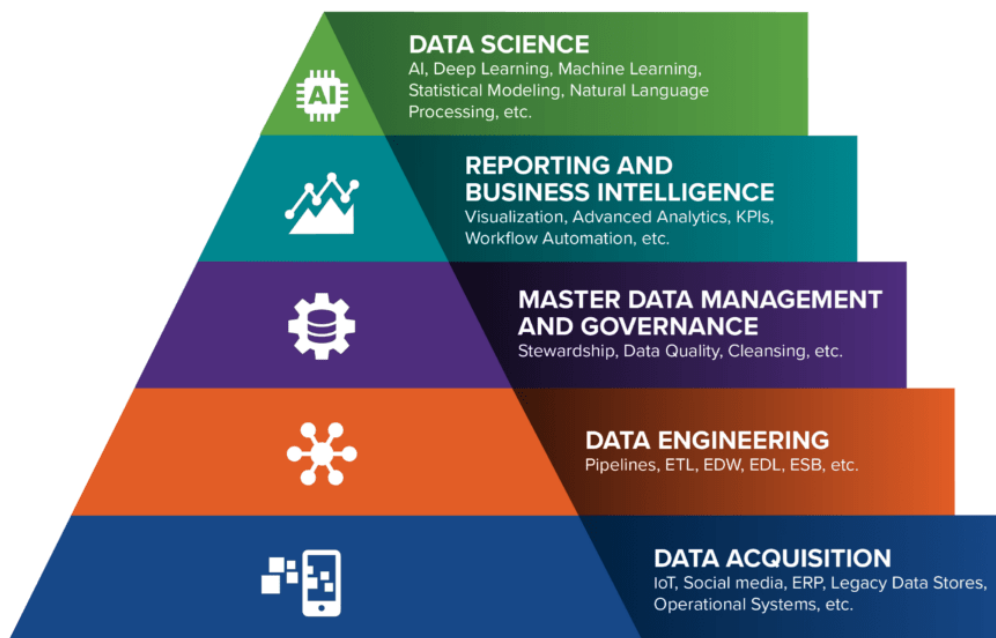
Data Lake e Data Warehouse non sono soluzioni antagoniste, semplicemente sono nate per scopi diversi.

Sarebbe opportuno cercare di creare una sinergia fra le due soluzioni: l'integrazione fra tutti i dati aziendali di qualsiasi livello e di qualsiasi formato (non solo i dati strutturati di un ERP, ma anche documenti video, immagini, ecc.) permetterebbe di catturare tutta **la conoscenza dell'azienda.**

Consigliamo di **evitare di lasciar diventare i Data Lake delle "paludi informative", utilizzati al massimo come aree prototipali dagli specialisti del dato** (che possono testare nuovi algoritmi per analizzare i dati), ma piuttosto pensare ad un'integrazione piena con le tradizionali architetture di Business Intelligence (Data Warehouse).

Infine parlando di tecnologie che collaborano non possiamo non ricordare che la **Cyber Security è assolutamente complementare ai Big Data analytics.** Vedi ad esempio i temi di **Identity and Access Management e Data Protection (GDPR)**

Metodologie e tecniche (La piramide operativa)



Intelligenza artificiale, Machine Learning e l'automazione delle analisi: **il cervello espanso del processo**

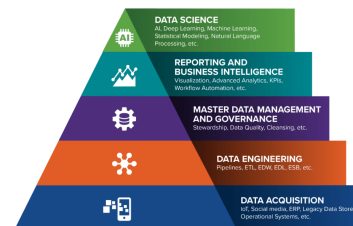
Le visualizzazioni e le analisi avanzate: **il volto del processo**

La qualità e la gestione dei datasets, finalizzato al mantenimento degli standard: **Il corpo del processo**

Progettare, installare, testare e mantenere i sistemi di gestione dei dati: **Lo scheletro del processo**

I sistemi o i metodi di acquisizione dei dati: **La base di tutto il processo**

Metodologie e tecniche (esempio)



1. Acquisizione dati



Raccogliamo dati (esempio qualità dell'aria) dal campo grazie a sistemi IoT e centraline ARPA

2. Ingegnerizzazione dati



Estraiamo, trasformiamo e carichiamo i dati in un architettura di sintesi (esempio database con informazioni sull'aria)

3. Mantenimento dati



Uniformiamo, puliamo e gestiamo coerentemente la semantica dei dati legati alla qualità dell'aria

4. Analisi e visualizzazione



Qui intervengono le metodologie e tecniche di analisi (data mining) e visualizzazione (data visualization) per dare valore ai dati e trasformarli in informazioni



5. Analisi avanzata

Sfruttiamo tecniche AI per automatizzare le analisi e scoprire correlazioni non evidenti al livello precedente

Metodologie e tecniche (data mining)

Il data mining è l'insieme di tecniche e metodologie utilizzate per estrarre di informazioni utili da grandi quantità di dati (es. banche dati, Data Warehouse, ecc.), attraverso metodi automatici o semi-automatici per utilizzo scientifico, industriale o operativo.



Il data mining **viene utilizzato per cercare correlazioni tra più variabili relativamente a un fenomeno**; ad esempio, conoscendo il comportamento medio dei clienti di una azienda, si cerca di prevedere quanto spenderà il cliente medio in futuro.

In sostanza il data mining è l'analisi, da un punto di vista matematico, eseguita su banche dati, preceduta tipicamente da altre fasi di preparazione, trasformazione o filtraggio dei dati come il data cleaning. Ci sono quindi due scopi:

- **estrarre, con tecniche analitiche all'avanguardia, informazioni implicite, nascoste, da dati già strutturati**, per renderle disponibile e direttamente utilizzabili
- **esplorare ed analizzare, in modo automatico o semiautomatico, grandi quantità di dati al fine di scoprire pattern** (schemi o regolarità) significativi

Metodologie e tecniche (data visualization)

La Data visualization è l'attività di comunicazione realizzata tramite la **proiezione di dati in una forma grafica strutturata**.

Questa proiezione è uno degli strumenti dell'analisi dati e della scienza dei dati e viene usata per **incrementare l'usabilità dei dati, rendendoli più accessibili, comprensibili e utilizzabili**.



La rappresentazione grafica dei dati utilizza vari elementi visivi come grafici statistici, diagrammi, tabelle, mappe, ecc. **Le dashboard o cruscotti sono gli strumenti che incorporano tutti questi elementi** permettendo di analizzare le informazioni in maniera rapida e intuitiva.

Le organizzazioni possono sfruttare queste informazioni per valutare e stimare il proprio contesto e il mercato di interesse.

Le aziende dovrebbero utilizzare le informazioni raccolte attraverso una strategia di **business intelligence per incrementare il loro vantaggio competitivo**.

Tipologie di analytics

Descriptive analytics

Le analisi descrittive **sintetizzano o descrivono i dati grezzi e creano qualcosa che è interpretabile dagli esseri umani.**

Vengono analizzati gli eventi passati, dove per eventi passati ci si riferisce a qualsiasi punto del tempo in cui si è verificato un evento. **Le analisi descrittive sono utili in quanto consentono alle organizzazioni di imparare dai comportamenti passati e aiutano a capire come potrebbero influenzare i risultati futuri.**

Predictive analytics

Le analisi predittive riguardano la **comprensione del futuro**. Essa fornisce alle organizzazioni informazioni intuitive basate su dati e **stima circa la probabilità di un futuro risultato**. Ma è altrettanto importante ricordare che nessun algoritmo statistico può “prevedere” il futuro con la precisione del 100%.

Prescriptive analytics

Questo campo relativamente nuovo dell'analisi Big Data, **facilita gli utenti a “prescrivere” diverse azioni possibili** per guidare l'attività verso una soluzione. **Tenta di quantificare l'effetto delle future decisioni per consigliare i possibili risultati** prima che esse siano effettivamente adottate. **L'analisi prescrittiva non solo prevede che cosa accadrà, ma spiega perché accadrà e fornisce raccomandazioni** in merito alle azioni per sfruttare queste previsioni.

Spunti e considerazioni

Quando la capacità di elaborazione interna all'impresa non è sufficiente o le elaborazioni sono troppo lunghe è **crescente la richiesta di sistemi di calcolo di tipo High Performance Computing (HPC)**, ossia di sistemi di calcolo che garantiscono prestazioni elevate attraverso tecniche di elaborazione in calcolo parallelo.

In Emilia-Romagna si concentra **il 70% di capacità di calcolo e di storage di tutta Italia (DATA VALLEY)**. Presso gli HPC sono **noleggiabili slot di calcolo come anche consulenza per analisi dati**.

Infrastrutture e servizi HPC sempre più rivolte anche alle PMI. Per analisi, simulazione, modellazione, prototipazione e progettazione che consentono elaborazioni più veloci (quindi soluzioni) a costi inferiori.

Inoltre oggi **sono disponibili alcune piattaforme open source (Hadoop, Spark, Kafka) che permettono di gestire grandi moli di dati su piccole reti di normali desktop** o su server virtuali noleggiati da fornitori di servizi di cloud computing (Amazon, Azure, Google).

CINECA

associazione
BigData
from volume to value

**Data Valley
Emilia-Romagna**

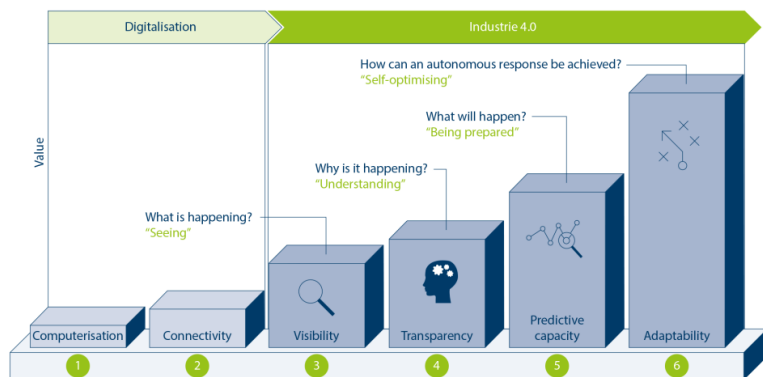


Spunti e considerazioni

Attenzione a non confondere i servizi web che consentono di analizzare dettagliatamente le statistiche sui visitatori di un sito web (Google Analytics), come un processo di big data analytics.

Infatti si tratta di **ANALYSIS** -> Cosa è accaduto e come. Si limita alle statistiche. Il focus è sul passato.

Mentre gli **ANALYTICS** -> sono la somma di operazioni che estraggono informazioni dai dati, tramite correlazioni che portino a un valore aggiunto per l'azienda. Si rivolge a scenari futuri.



I data analytics abilitano la Trasparenza:

l'azienda utilizza tecnologie per l'analisi dei dati finalizzate a comprendere le interazioni presenti tra gli elementi dell'ombra digitale.

Tutti gli ambiti possono essere interessati

Agricoltura



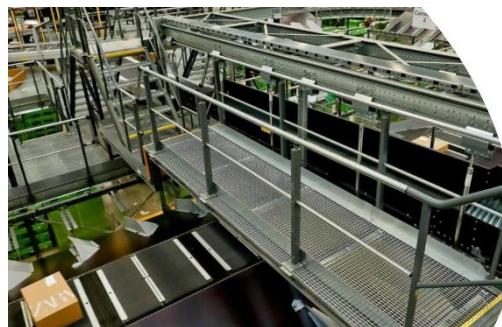
Commercio



Artigianato



Industria



Agricoltura

Il quaderno di campagna 4.0

La sfida:

Rendere il quaderno di campagna e il registro dei trattamenti degli strumenti utili a valutare le scelte operative che si fanno, i protocolli seguiti e le tecniche utilizzate, confrontando a fine stagione produttiva i risultati ottenuti e poter pianificare le eventuali azioni correttive in maniera data-driven.

La soluzione:

Strumenti digitali in grado di supportare le imprese nella programmazione quali-quantitativa delle produzioni e avere la possibilità di prevedere, documentare e gestire, le diverse fasi dei processi attraverso una efficiente gestione dei dati. Per esempio, i piani di fertilizzazione e trattamento per le colture vengono calcolati sulla base di banche dati dedicate e alle informazioni provenienti dal campo. La soluzione controlla gli apporti massimi di azoto, fosforo e potassio, segnalando eventuali superamenti e predisponendo poi le stampe previste dalla normativa vigente (D.Lgs. 150/2012).

I benefici:

- Limitare le attività di compilazione manuale del quaderno di campagna
- Possibilità di pianificazione strategica guidata dai dati storici
- Snellimento e efficientamento delle procedure richieste a norma di legge



Artigianato

Attività di ottimizzazione del design di prodotto

La sfida:

Ottimizzare il design delle tavole da surf

La soluzione:

Il design delle tavole da surf rappresenta un raffinato mix di arte e scienza. Oggi le tecnologie digitali e i modelli computazionali possono rappresentare uno strumento in grado di supportare e migliorare il design delle tavole da surf affiancandosi all'esperienza ed alla tradizione.

L'applicazione della fluidodinamica computazionale (CFD) al design delle tavole da surf attraverso la simulazione del flusso turbolento a superficie libera attorno ad una tavola può aiutare gli artigiani "shapers" a innovare e ottimizzare il proprio prodotto in maniera data-driven.

I benefici:

- Ottimizzazione delle performance del prodotto
- Riduzione dei costi per i materiali
- Innovazione di prodotto



Commercio

Revisioni di vendita per l'ottimizzazione dei magazzini

La sfida:

Ottimizzare le merci a magazzino in base alle previsioni di vendita.

La soluzione:

Grazie all'analisi delle abitudini di acquisto dei clienti e al valore derivante dalla mole di dati relativa agli acquisti in generale (es. scontrini) è possibile per le aziende attivare strategie data-driven di gestione ottimizzata dei magazzini (compro solo quello che veramente mi serve in base alle abitudini di acquisto), programmazioni puntuali di riordino e gestione dell'obsolescenza delle merci, abilitando così modelli di business che valorizzino il fine vita dei prodotti invenduti.

I benefici:

- Riduzione costi di magazzino
- Proposta di valore orientata al cliente, monitorando le sue abitudini
- Aiuta a programmare le migliori strategie di riordino, qualità e obsolescenza delle merci



Industria

Ottimizzare il layout di fabbrica

La sfida:

Migliorare la posizione e l'operatività dei macchinari, dei magazzini e delle aree di lavoro.

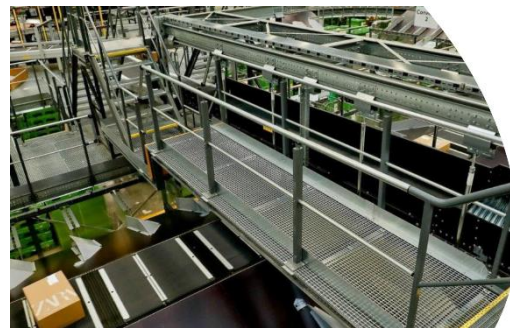
La soluzione:

Grazie a un software in grado di simulare e analizzare il comportamento di un processo è possibile far emergere le criticità, eventuali colli di bottiglia e accrescere il know-how del processo e il grado di conoscenza del management.

Sulla base dei risultati della simulazione è possibile ridurre le distanze tra le stazioni di lavoro e il magazzino, collocare diversamente le linee produttive e le aree di lavoro, relazionandole ai flussi di magazzino su cui possono verificarsi colli di bottiglia e riuscendo anche a ridurre i tempi morti degli operatori.

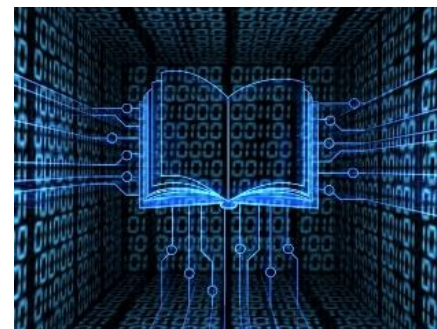
I benefici:

- Ottimizzazione del processo produttivo
- Riduzione dei costi di produzione e di magazzino
- Modellazione di sistemi gestionali e operativi complessi



Il focus: Gestione documentale e Archivio 4.0

I processi di gestione e conservazione degli archivi cartacei sono caratterizzati da costi e tempi di ricerca elevati, da un forte impatto ambientale e da un alto margine di errore e smarrimento. Questo è un **ambito in cui le tecnologie digitali possono supportare le aziende nel far evolvere i propri processi**



Per creare un archivio digitale è necessario innanzitutto un'appropriate analisi dell'archivio cartaceo per individuare la documentazione da scartare e quella da conservare, seguita dalla **definizione della modalità operativa ottimale** per garantire un **accesso agile, sicuro e completo ai documenti digitali**.

Immaginate infine un sistema digitale capace di fornire **automaticamente le etichette utili a catalogare un documento o un set di fotografie**, la catalogazione e la ricerca documentale avverrebbe in maniera più agile e snellirebbe i processi di gestione e consultazione dell'archivio.

Non è fantascienza. È una realtà legata all'uso di tecnologie del paradigma 4.0, quali **Intelligenza Artificiale** e **Big Data Analytics** che portano valore aggiunto nei processi di gestione e tracciabilità, abilitando la creazione dell'archivio 4.0.

L'archivio 4.0 (l'esperienza di NDM)

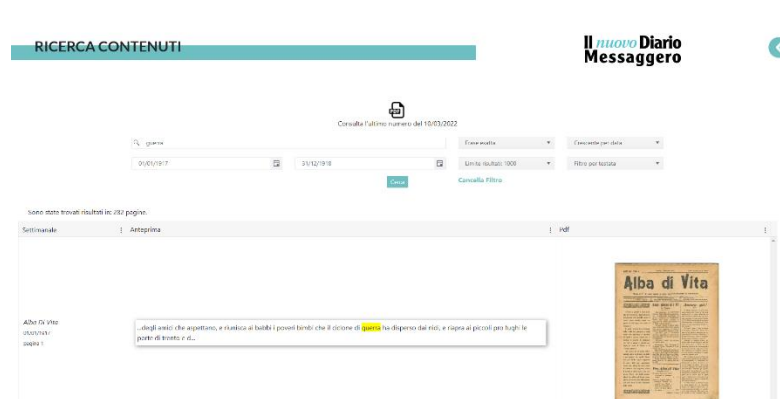
Il *nuovo* Diario Messaggero

Il Nuovo Diario Messaggero è la testata giornalistica della diocesi di Imola. **Nel 2019 ha attivato un percorso di digitalizzazione dell'archivio storico del giornale.**

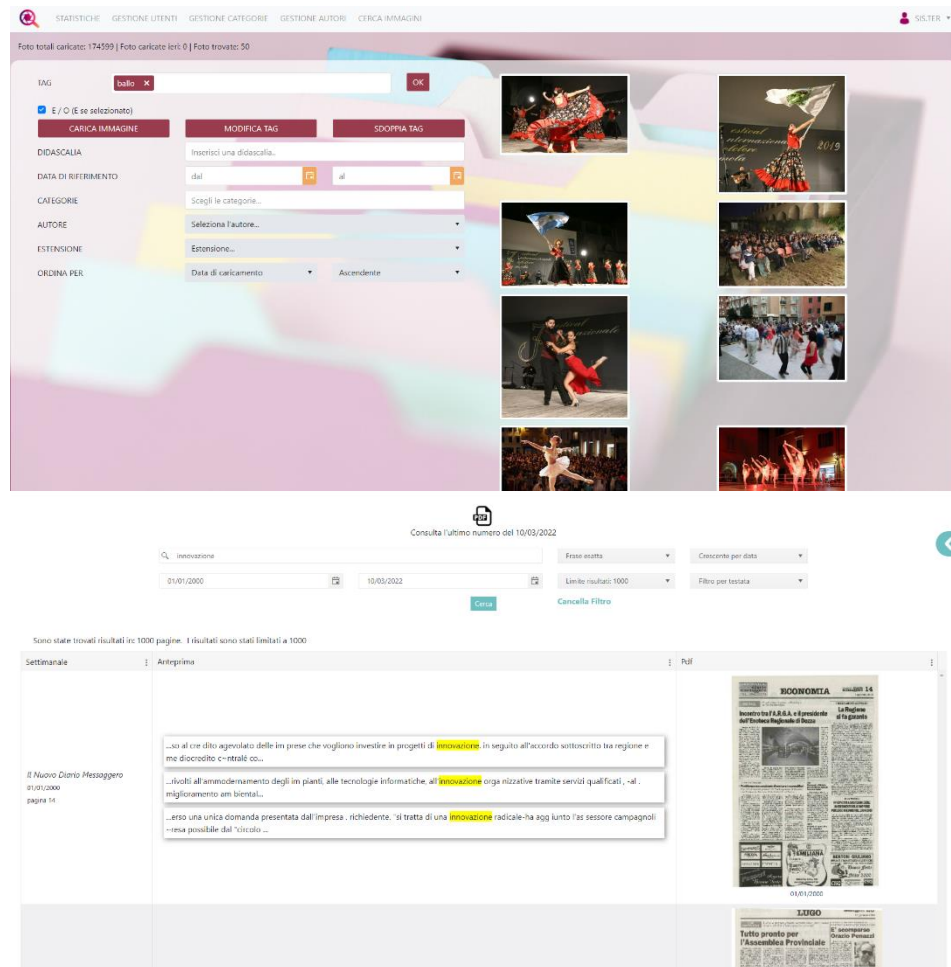
L'archivio storico digitale mette a disposizione degli utenti ogni pagina mandata in stampa dal 1892 (anno in cui è stato pubblicato il primo giornale) ad oggi. Oltre a questo anche tutti i bollettini diocesani dal 1914 ad oggi. Un vero e proprio patrimonio sociale e culturale della memoria storica non solo della città di Imola.

Il servizio web scaturito da questo progetto offre agli abbonati e a tutta la città imolese più di 125 anni di storia a portata di click.

La soluzione sviluppata da Sis.Ter srl ha reso consultabili tutte le pagine mandate in stampa dal 1892. Inoltre è stato sviluppato un **archivio 4.0 delle immagini presenti nelle pagine, fino a ospitare circa 175.000 foto.**



L'archivio 4.0 (l'esperienza di NDM)



Dopo aver catalogato e inserito nel sistema SW tutta la mole di dati relativi alle pubblicazioni del giornale. **Dati che sono filtrabili e ricercabili rapidamente grazie a un sistema di Tags (metadati).**

Grazie all'utilizzo di soluzioni di **Intelligenza Artificiale e Machine Learning (Amazon Image Rekognition)** abbiamo abilitato la ricerca automatica e la catalogazione tematica del corpus fotografico del giornale imolese.

Semplificando di molto il lavoro della redazione e abilitando un servizio nuovo.

Opportunità derivanti della tecnologia

Oggi non approcciare i **Big Data Analytics** equivale a perdere opportunità, in casi estremi equivale ad essere esclusi da nuovi mercati o addirittura da quelli in cui già si opera.

Le imprese che hanno già messo in atto **progetti che sfruttano i Big Data Analytics** lo hanno fatto per raggiungere questi obiettivi:

- **design di prodotto migliorato:** grazie ai dati di comportamento e funzionamento di un prodotto è possibile simulare scenari migliorativi
- **efficienza gestionale e produttiva:** ad esempio **riduzioni di tempi di inattività di un processo;**
- **opportunità di business:** le analisi dati abilitano lo sviluppo di nuove opportunità a valore aggiunto;
- **migliore customer experience:** legata alle analisi dei comportamenti dei clienti e alla proposta di valore data-driven;
- **ottimizzazione della agilità:** con la creazione di processi decisionali più flessibili e data-driven.

Sfide della tecnologia e approcci da evitare



Una sfida cruciale dei prossimi anni sia per le imprese che per le istituzioni è **lo sfruttamento di una mole di informazioni derivanti dai Big Data per una radicale revisione dei metodi di decisione ed elaborazione delle linee di azione.**

Tuttavia senza competenze tecniche e strumenti adeguati di analisi e gestione, i Big Data rischiano di restare solamente numeri interessanti, ma non in grado di divenire supporto significativo alle strategie d'impresa. Le aziende hanno bisogno di **investire in formazione per ampliare le competenze** delle proprie risorse, e di focalizzarsi sul costruire skills adeguate.

Cultura del dato, strumenti e metodologie innovative, progetti di Analytics e, soprattutto, grande competenza nel Data Management. Un consiglio per le aziende che vogliono davvero avere successo? **Essere sempre più data-driven.**

Attenzione infine alla qualità e alla pulizia dei dati che analizziamo, perché i computer elaborano in modo acritico anche un insieme di dati in entrata palesemente insensati o errati (*garbage in*) producendo, a loro volta, un risultato insensato (*garbage out*). Questo vale doppiamente nel caso si sfruttino tecniche di apprendimento automatico e intelligenza artificiale.



Rubrica sui finanziamenti per progetti 4.0

La nuova **Strategia S3** della regione ha individuato **8 aree di specializzazione strategica**. La Strategia S3 2021-2027 è stata definita con un **percorso partecipato** che ha coinvolto gli stakeholder e la comunità regionale dell'ecosistema della ricerca e dell'innovazione. Sono stati individuati **15 ambiti tematici**, derivanti dall'incrocio fra i sistemi produttivi regionali e le sfide della Politica di coesione UE 21-27

Regione ER	Opportunità varie (portale da monitorare spesso)
ART-ER	Programmi di Open Innovation e ricerca
ER-STARTUP	Bandi specialmente dedicati alle Start Up
Unioncamere	Opportunità orientate all'internazionalizzazione

Rubrica sui finanziamenti per progetti 4.0

BOLLETTINO MENSILE CCIAA-BO

www.bit.ly/PID_BO_Contributi

Unioncamere Emilia Romagna

<https://www.ucer.camcom.it/Internazionalizzazione/bandi>

Emilia Romagna finanziamenti per imprese

<https://imprese.regione.emilia-romagna.it/Finanziamenti/finanziamenti-aperti>

ART-ER contributi economici

<https://www.art-er.it/bandi-e-avvisi/contributi-e-vantaggi-economici/>

Emilia Romagna START- UP bandi

<https://www.emiliaromagnastartup.it/it/bandi>

S3 Smart Specialization Strategy 2021-2027

<https://fesr.regione.emilia-romagna.it/s3/2021-2027>

GRAZIE PER L'ATTENZIONE

Q&A SESSION

Vi diamo appuntamento alla prossima pillola formativa il
13/04/2022

Parleremo di **Integrazione verticale e orizzontale e
Sistemi gestionali**