

POLITECNICO DI TORINO

Dipartimento di Ingegneria Gestionale e della Produzione

Corso di Laurea Magistrale in Ingegneria Gestionale



TESI DI LAUREA

Data Management: implicazioni organizzative e gestionali per
imprese nel settore manifatturiero

Il caso Brugnago S.R.L.

Anno accademico 2020 – 2021

Studente:

DAVIDE TARABBO

Relatore:

Prof. EMILIO PAOLUCCI

Matricola: 267788

Indice

Indice.....	2
Introduzione.....	4
1 I DATI.....	7
1.1 Definizione.....	7
1.2 Crescita dei volumi, delle modalità di gestione, evoluzione delle tecnologie	8
1.3 Valore dei dati	10
1.3.1 Il ritorno aziendale sui dati	11
1.4 Come estrarre valore dai dati	13
2 IL DATA MANAGEMENT.....	19
2.1 Definizione di Data Management.....	19
2.2 Knowledge areas	21
2.2.1 Data Governance	22
2.2.2 Data Architecture	24
2.2.3 Data Modeling and Design	26
2.2.4 Data Storage and Operations	28
2.2.5 Document and content management	29
2.2.6 Data Security.....	30
2.2.7 Data Integration and Interoperability	32
2.2.8 Data Warehousing and Business Intelligence.....	35
2.2.9 Data Quality	38
2.2.10 Metadata	39
2.2.11 Reference e Master Data.....	40
2.3 Come esercitare il Data Management in pratica.....	41
3 DATI E MERCATI.....	42
3.1 Data Management nel settore finanziario	42
3.2 Data Management nel settore energetico	45
3.3 Data Management nel settore manifatturiero.....	48
4 IL CASO REALE DELL'AZIENDA BRUGNAGO S.R.L.....	52
4.1 L'azienda	52
4.2 I dati dell'azienda.....	53
4.3 Il problema.....	55
4.4 Proposta di soluzione	55

4.5	Soluzione 1: automatizzazione dei processi di raccolta ed integrazione dati per il monitoraggio dell'indice Overall Equipment Effectiveness	56
4.5.1	L'indice Overall Equipment Effectiveness OEE	56
4.5.2	Fonti di dati.....	58
4.5.3	Situazione attuale per il monitoraggio dell'indice OEE	61
4.5.4	Problemi legati alla situazione attuale	61
4.5.5	Implementazione dell'integrazione automatizzata.....	62
4.5.6	Situazione dopo l'implementazione	63
4.5.7	Benefici	63
4.6	Soluzione 2: sviluppo di un modello di gestione dati e di un glossario di business	65
4.6.1	Fasi di sviluppo del modello	66
4.6.2	Descrizione modello	67
4.6.3	Benefici	70
4.7	Le due soluzioni graficamente.....	72
4.8	Analisi dei costi	73
4.9	Analisi del cambiamento culturale in azienda.....	75
4.10	Benefici e creazione del valore per l'azienda	76
4.11	Condizioni necessarie per procedere con investimenti	78
	Conclusioni	80
	Fonti bibliografiche e sitografia	83
	Ringraziamenti	85

Introduzione

Al giorno d'oggi si parla sempre di più di impresa connessa, un termine che non riguarda più solo il processo produttivo, bensì abbraccia tutte le funzioni aziendali, al di fuori delle mura dei singoli reparti, ma anche al di fuori di una singola azienda, snodandosi lungo tutta la supply chain. In questo caso il passaggio fondamentale nell'evoluzione tecnologica consiste nell'uso di tecnologie avanzate di informazione e comunicazione e di analisi avanzate dei dati. Gli studi compiuti fino ad ora, mostrano come la crescente digitalizzazione all'interno delle imprese manifatturiere comporti la creazione di una quantità sempre maggiore di dati, ponendo così le aziende di fronte a nuove sfide. È necessario, infatti, che le pratiche attuali per l'archiviazione e la gestione dei dati siano in grado di resistere all'evoluzione delle dimensioni e della velocità dei dati forniti. La letteratura evidenzia inoltre come molte aziende non abbiano ancora sviluppato soluzioni organizzative e tecnologiche, necessarie per una corretta gestione e utilizzo dei dati interni. Trattare i dati come una risorsa strategica può creare un vantaggio competitivo se tali dati sono standardizzati e interoperabili in tutta l'organizzazione. In alcuni casi, se i dati non vengono gestiti correttamente, pur essendo disponibili, possono portare con sé problemi di affidabilità.

Un recente sondaggio condotto da Boston Consulting Group (BCG) su oltre 1.300 dirigenti di aziende nel settore manifatturiero, ha rilevato che il 72% delle aziende considera l'analisi di dati e la disciplina di Data Management più importante ora di quanto non lo fosse tre anni fa. I produttori che hanno partecipato all'indagine hanno citato varie sfide che hanno ostacolato i loro sforzi per scalare e implementare ulteriormente soluzioni di Data Management all'interno dei loro stabilimenti. Tra i problemi principali sono stati evidenziati: la mancanza di competenze e strumenti per una corretta gestione e analisi dei dati, ma anche una definizione non chiara di ruoli e responsabilità. Nasce quindi la necessità da parte delle imprese manifatturiere di sviluppare nuove competenze tecnologiche e di capitale umano, permettendo così di considerare le imprese 4.0 un vero e proprio incubatore di skills e capabilities legate alla digitalizzazione.

Alla luce di queste ricerche già svolte nel corso degli anni, con il lavoro di tesi ho voluto esplorare alcune delle tematiche sopra presentate. L'elaborato si concentra sull'introduzione di procedure di Data Management all'interno di imprese del settore manifatturiero. Il lavoro si è incentrato sulle sfide che un'azienda di questo settore deve affrontare a fronte della digitalizzazione e della gestione

dei dati che ogni giorno vengono prodotti internamente. Le domande e ipotesi sulle quali si è concentrata la tesi sono state principalmente: quali sono gli investimenti necessari da compiere in un'impresa manifatturiera per introdurre procedure di Data Management? Quali competenze è necessario sviluppare per rispondere ad un inevitabile cambiamento culturale aziendale? Più in generale, quindi, quali sono e quanto sono grandi gli ostacoli che un'azienda deve affrontare di fronte al cambiamento organizzativo, volendo introdurre strumenti e procedure di Data Management al suo interno? Quali sono i possibili benefici che derivano da un uso migliore dei dati aziendali?

Come mostrato dal primo capitolo, il tema è più che attuale, dal momento che il numero di dati presenti sul nostro pianeta, in qualunque settore considerato, è in continuo aumento. Per comprendere meglio la disciplina di Data Management, è stato descritto nel secondo capitolo il framework proposto dall'associazione DAMA International. In questo modo la disciplina diventa accessibile ad un pubblico sempre più ampio, permettendo così, attraverso procedure ben definite, di utilizzare pratiche per la gestione dei dati. L'analisi prosegue nel terzo capitolo con una valutazione del Data Management nei principali settori: finanziario, energetico e manifatturiero. Attraverso l'analisi di questo capitolo è stato possibile mappare, per aziende in diversi settori, l'attuale situazione di gestione dei dati a livello internazionale. Per rispondere alle domande di ricerca sopra citate, nel quarto capitolo è stato analizzato il caso d'uso della manifattura Brugnago S.R.L. All'interno dell'azienda, che presenta un livello medio - alto di digitalizzazione, è stato sviluppato un progetto di Data Management, in collaborazione con l'azienda Irion S.R.L. Attraverso il caso studio dell'azienda Brugnago, è stato possibile analizzare la quantità e le caratteristiche principali dei dati che vengono creati e utilizzati all'interno di una PMI del settore manifatturiero. L'attività di consulenza di Irion per offrire soluzioni di Data Management ha consentito di proporre due soluzioni per la manifattura Brugnago. Con la prima soluzione è stato rivisto un flusso di dati interno all'azienda, proponendo un'integrazione dei dati automatizzata utile a realizzare il calcolo dell'indice di produttività Overall Equipment Effectiveness (OEE). Con la seconda soluzione è stato creato un modello di gestione dati attraverso gli strumenti dell'azienda Irion. Per entrambe le soluzioni sono stati valutati i benefici potenziali derivanti dalla loro adozione. L'analisi si è poi concentrata sugli investimenti necessari per l'introduzione di procedure, strumenti e competenze di Data Management all'interno dell'azienda. In conclusione sono stati considerati i benefici e la possibile creazione di valore per l'azienda a seguito di una migliore governance e utilizzo dei dati.

Con la crescente digitalizzazione e con l'aumento di dati prodotti le aziende devono reagire e trovare nuove soluzioni per la gestione dei dati aziendali. Di fronte alla nuova rivoluzione industriale 4.0 e al concetto sempre più attuale di impresa connessa, investimenti in hardware, software e competenze per supportare la digitalizzazione e l'innovazione sono fondamentali per evitare che l'evoluzione tecnologica sia disruptive. Questo periodo di incertezza e cambiamento tecnologico può far nascere importanti opportunità per la creazione di valore e vantaggio competitivo sul mercato se le nuove competenze vengono sfruttate dalle imprese in modo consapevole, puntando sui giusti obiettivi di business con una visione chiara di lungo periodo.

1 I DATI

1.1 Definizione

Il New Oxford American Dictionary definisce i dati come “fatti e statistiche raccolti insieme per l’analisi”. L’American Society for Quality definisce i dati come “Serie di fattori raccolti”. In generale oggi siamo in grado di acquisire moltissime informazioni attraverso i dati. Lo sviluppo tecnologico sta rendendo il mondo sempre più “Data Driven”, dal momento che la diffusione di dispositivi elettronici in grado di raccogliere, produrre, gestire e scambiare dati è sempre più frequente. Alcuni esempi possono essere smartphone, computer, elettrodomestici “intelligenti”, smart meter ed in generale IoT. Questi fatti e queste informazioni possono essere aggregate, analizzate e utilizzate per ottenere profitto e vantaggio competitivo da parte delle aziende.

Molto spesso le persone pensano che i dati possano essere considerati una forma di verità assoluta e che concordino tra di loro. In realtà i dati sono un mezzo di rappresentazione, e rappresentano quindi cose differenti tra di loro. I dati sono sia un’interpretazione di oggetti che essi rappresentano sia un oggetto che deve essere interpretato (Sebastian-Coleman, 2013). Questa affermazione introduce un concetto importante, ovvero che i dati abbiano bisogno di un contesto per essere significativi. Questo contesto può essere visto come un sistema di rappresentazione dei dati, che comprende quindi un dizionario comune e una serie di relazioni tra i vari componenti. È necessaria quindi una conoscenza delle convenzioni del sistema, in maniera tale da poter essere in grado di interpretare i dati al suo interno. A proposito di questo, l’International Standards Organization (ISO) definisce i dati come rappresentazione di informazioni reinterpretabile in maniera formalizzata adatta alla comunicazione, interpretazione o elaborazione. Questa definizione, da un lato sottolinea la natura elettronica dei dati, mentre dall’altro lato, presume che i dati necessitino di standard, dal momento che vengono gestiti da sistemi di information technology.

Le persone tendono a fare scelte diverse su come rappresentare i concetti, in particolare, generano modi differenti per rappresentare concetti identici. Da queste scelte, i dati assumono forme diverse, che portano alla necessità di introdurre elementi di Data Management che permettano di comprendere ed utilizzare in maniera più sapiente i dati. Anche all’interno di un’organizzazione esistono molti modi per rappresentare la stessa idea, e tra più organizzazioni il problema della

molteplicità aumenta ulteriormente. Per questo motivo nasce la necessità di standard a livello industriale che possano conferire maggiore coerenza ai dati.

Molto spesso nel corso degli anni si è discusso sulla relazione tra dati ed informazioni. I due concetti sono collegati e dipendenti l'uno dall'altro. A livello organizzativo è sempre e comunque meglio tracciare una linea tra informazioni e dati per scopi di comunicazione chiara sui requisiti e sulle aspettative dei vari utilizzi da parte dei diversi stakeholder che ne necessitano. Ad esempio, se si volesse fare un resoconto sulle vendite dell'ultimo trimestre (informazioni) sarebbero necessari dati provenienti dal Data Warehouse aziendale. D'altro canto, i risultati sotto forma di dati potranno essere utilizzati per generare misure di performance per i successivi trimestri, ovvero un flusso di informazioni. Da qui si può dedurre uno dei principi cardine del Data Management: sia i dati che le informazioni devono essere gestiti. Inoltre, entrambi saranno di migliore qualità se gestiti insieme in base ad usi e requisiti dei clienti.

Da sempre le organizzazioni utilizzano i dati per la gestione ed il monitoraggio dei propri processi operativi, delle proprie performance e per l'analisi del mercato, ma l'evoluzione delle tecnologie dell'informazione ha reso disponibili modalità di impiego dei dati sempre più potenti, robuste e sofisticate. Questo fenomeno ha favorito una consapevolezza ed una attitudine delle organizzazioni e dei singoli sul valore delle informazioni. Questi cambiamenti sono risultati positivi, in quanto hanno permesso alle organizzazioni di utilizzare i dati in modi nuovi in modo da creare prodotti, condividere informazioni, migliorare l'efficienza aziendale, contribuire ad accrescere il valore d'impresa. Nonostante ciò, la rapidità della crescita tecnologica e con essa la relativa capacità dell'uomo di produrre, acquisire ed estrarre dati ha reso ancor più indispensabile la necessità di gestire i dati in modo efficace. Da qui nasce l'esigenza del Data Management e di procedure affidabili ed estensibili, in modo da trarre vantaggio dalla grande varietà di dati a disposizione senza esserne sommersi dal volume e dalla velocità.

1.2 Crescita dei volumi, delle modalità di gestione, evoluzione delle tecnologie

La necessità di promuovere una gestione efficiente dei dati tramite procedure mirate di Data Management è testimoniata dal fatto che, negli ultimi anni ci sia stata una crescita esponenziale dei dati a livello globale. Le stime degli anni precedenti prevedevano un incremento elevatissimo attorno al 2020 e a seguire negli anni. In particolare diverse fonti erano tutte ampiamente d'accordo sul fatto che la dimensione delle informazioni digitali sarebbe raddoppiata almeno ogni due anni,

tra il 2010 ed il 2020, tendendo poi ad una crescita esponenziale. I dati generati da uomini e macchine stanno crescendo ad una velocità dieci volte maggiore rispetto ai tradizionali dati di business. I dati macchina, come dati generati da sensori, dati che creano altri dati attraverso calcoli, previsioni, automazioni, metadati, transazioni, stanno crescendo ancora più rapidamente, ad un tasso di crescita di cinquanta volte. Per comprendere la reale quantità di dati che circonda l'uomo ai giorni nostri, è interessante uno studio di IDC, leader a livello mondiale in ricerche di mercato, consulenza in ambito IT ed innovazione digitale. Il datasphere globale di IDC misura la quantità di dati creati e consumati ogni anno nel mondo. Il rapporto tra dati univoci (creati e acquisiti) e dati replicati (copiati e consumati) è di circa 1:9, ma la tendenza è una migrazione lenta verso dati meno univoci e più replicati. Entro il 2024, IDC si aspetta che il rapporto diventerà di 1:10. "La crescita del Global DataSphere è guidata più dai dati che consumiamo e analizziamo che da ciò che creiamo", sono le parole del vice presidente di IDC's Global DataSphere, David Reinsel. Egli sostiene inoltre che, naturalmente i dati debbano essere creati prima di poter essere analizzati, ma il tasso di ricorsione dei dati, ovvero la velocità con cui gli stessi dati vengono elaborati di nuovo, continua a crescere in modo esponenziale facendo diminuire il DataSphere "unico" fino al 10% del DataSphere totale.

Alcuni aspetti chiave derivanti dalla previsione del Global DataSphere possono includere:

- La quantità di dati che verranno creati a partire dal 2020 nei prossimi tre anni sarà maggiore dei dati creati negli ultimi 30 anni, ed il mondo creerà più di tre volte i dati nei prossimi cinque anni rispetto ai cinque precedenti.
- La produttività/dati incorporati è la categoria in più rapida crescita per la creazione di dati con un CAGR (compound annual growth rate) del 40,3% per il periodo di previsione 2019-2024. Entro il 2024, i dati di intrattenimento rappresenteranno il 40% della Global DataSphere e i dati di produttività / incorporati saranno del 29%, in qualche modo bloccati dalle dinamiche COVID-19.
- I sensori vengono incorporati in qualsiasi oggetto. Il mondo che ci circonda è costituito da una grandissima quantità di sensori, dalle automobili agli elettrodomestici, dagli smartphones ai sistemi di illuminazione automatica ed in generale tutto il mondo di IoT. Questi generano dati che possono aiutare a contestualizzare i dati stessi. Questi dati insieme a quantità crescenti di metadati, stanno crescendo in modo aggressivo e presto raggiungeranno una quantità ingente, potendo addirittura superare gli altri tipi di dati.

- Dati di videosorveglianza e privacy e iniziative/mandati normativi continuano ad intersecarsi. Trovare il giusto equilibrio tra sicurezza, personalizzazione, efficienza e diritti individuali alla privacy sarà una delle grandi tensioni di questo decennio.

IDC inoltre predice che i dati a livello mondiale cresceranno entro il 2025 fino ad un numero di 175 zettabytes, ovvero 175 trilioni di gigabytes, attraverso un tasso di crescita annuale del 61%.

Le attività di Data Management richiedono competenze sia tecniche che di business. Dal momento che quasi tutta la totalità dei dati di oggi è memorizzata elettronicamente, le tecniche di Data Management sono fortemente influenzate dalla tecnologia. Per questo motivo, un ruolo sicuramente importante nella gestione ed utilizzo dei dati è rappresentato dall'evoluzione tecnologica che ha accompagnato il mondo del Data Management in questi anni. Un esempio può essere rappresentato dalla disciplina di Data Science. Questa infatti, esiste da molto tempo, ed è in generale un termine che viene utilizzato per definire "statistiche applicate". La business intelligence tradizionale prevede l'analisi di dati strutturati per descrivere tendenze passate, ed in alcuni casi, gli schemi di BI vengono utilizzati per predire comportamenti futuri, anche se non con grande accuratezza. In questo caso il cambiamento tecnologico risulta fondamentale. Fino a poco tempo fa, infatti, l'analisi approfondita di grandi quantità di dati era fortemente limitata dalla tecnologia. Con l'aumento della capacità di raccolta, tramite nuovi database e sistemi di integrazione, e di analisi dei dati stessi, i Data Scientist hanno integrato metodi da formule, statistiche, scienze informatiche, signal processing, modellazione probabilistica, pattern recognition, machine learning, modellazione dell'incertezza e visualizzazione dei dati per poter avere una miglior visione e predire i comportamenti basandosi proprio su serie di Big Data.

1.3 Valore dei dati

I dati, a differenza di asset fisici che possono essere toccati e dislocati e possono trovarsi in un solo posto alla volta, e, a differenza degli asset finanziari, che vengono contabilizzati in bilancio, non sono tangibili. I dati però sono diversi anche sotto altri punti di vista: sono durevoli, non si consumano, ma il loro valore varia lungo il loro ciclo di vita. Sono facili da copiare e trasportare. I dati sono dinamici e possono essere utilizzati per più scopi e contemporaneamente da più persone, rendendoli così una risorsa flessibile e con forti potenzialità per l'organizzazione, dal momento che nessun asset fisico possiede queste caratteristiche. È necessario procedere però con cautela, dal momento che spesso l'uso di dati genera a sua volta altri dati e di conseguenza le organizzazioni si

ritrovano a dover gestire un volume crescente di dati e relazioni tra set di dati stessi, accrescendo così sia il vantaggio dovuto ai dati a disposizione ma anche lo sforzo nel gestirli.

Queste considerazioni e differenze rispetto agli asset fisici, rende difficile assegnare un vero valore monetario ai dati e naturalmente, senza questo valore, è difficile quantificare in che modo un dato contribuisca al successo organizzativo. Oltre a nuovi vantaggi derivanti dal possesso e dall'uso di dati, nascono anche nuove problematiche che riguardano il Data Management, come la definizione della proprietà dei dati, effettuare l'inventario dei dati a disposizione di un'organizzazione, la protezione contro un uso scorretto dei dati, definizione di rischi associati all'uso non corretto dei dati e la definizione e applicazione degli standard per la qualità dei dati.

Nonostante le difficoltà nella valutazione del valore effettivo dei dati, molte organizzazioni riconoscono che i dati abbiano un valore. Il primo aspetto fondamentale è che i dati di un'organizzazione sono di per sé unici, la loro distruzione o perdita potrebbe avere conseguenze rilevanti per i risultati aziendali.

I dati e le informazioni, come già visto, sono strettamente collegati tra di loro, sono fondamentali per gli affari dell'organizzazione. Molte operazioni di business coinvolgono uno scambio di informazioni. La maggior parte delle informazioni si muove elettronicamente, generando una traccia che può servire per monitorare i movimenti dei dati, generare uno storico e poi svolgere analisi utili sui movimenti dei dati.

1.3.1 Il ritorno aziendale sui dati

Uno dei temi fondamentali che circonda il mondo dei dati e che coinvolge quindi anche direttamente quello del Data Management, è la capacità di quantificare il ritorno sugli investimenti fatti per aumentare il valore ottenuto dai dati. Nel corso degli anni i dati sono stati definiti come il nuovo petrolio, ma molto spesso invece anche come vecchio petrolio e inutilizzabile. Il problema infatti, è che nel corso degli anni non è stato chiaramente visibile un andamento nei flussi di cassa direttamente imputabile ai dati, tendendo così ad incriminarli. Ma la vera soluzione sta infatti nel mettere in comunicazione le capacità che consentono la comprensione ed il valore aziendale risultante.

Nel corso di un sondaggio rivolto a dirigenti di diverse compagnie in vari settori, otto dei dieci rispondenti ha affermato che fosse importante misurare e riportare i risultati di business derivanti

dai loro dati e dagli investimenti in analytics. Ma molte organizzazioni hanno difficoltà a raggiungere questo obiettivo, come confermato dal fatto che soltanto un intervistato su 10 riesca a misurare e riportare in modo efficace l'impatto sul business dell'analisi dei dati e degli investimenti. Il grosso problema è che il collegamento tra dati, analytics e risultati di business può essere un tema molto astratto, ma misurare e riferire gli obiettivi ottenuti con piccole vittorie può illustrare cosa sia possibile fare su una scala più ampia. Questa teoria può essere dimostrata dall'atteggiamento dei leader di aziende e compagnie e dalle loro decisioni, che stanno prendendo per differenziarsi e dare una spinta al loro business in modo da avere successo dove gli altri invece lottano. Infatti, le loro scelte sull'adozione e gestione di strumenti multi-cloud, gli investimenti nelle capabilities necessarie, la democratizzazione dei dati e le analisi all'interno di un'organizzazione sembrano coincidere con reali benefit di business.

Una decisione strategica sul lancio di un nuovo prodotto o sull'ingresso in un nuovo segmento di mercato che garantisca un payoff nel lungo periodo, necessita di una combinazione di analisi qualitative e quantitative per poter valutare la qualità della decisione presa. Proprio i dati risultano fondamentali in questo compito, i quali non possono essere quindi valutati in cassa sul momento per dare valore all'azienda, ma permettono di avere una maggiore precisione nelle decisioni strategiche aziendali e garantire quindi risultati economici migliori e più affidabili. Da decisioni migliori infatti nascono risultati migliori. Durante un sondaggio, i maggiori leader aziendali lo hanno confermato, dicendo inoltre che non solo hanno riscontrato una maggiore efficienza operativa rispetto agli anni precedenti, ma anche un passaggio significativo verso metriche di performance più strategiche. Inoltre, anche la soddisfazione degli impiegati è migliorata come anche la loro lealtà nei confronti dei loro leader e aziende. Infine, è interessante notare come le organizzazioni che siano più mature nelle loro analisi dei dati possano aumentare la velocità dei cicli di innovazione del business, creando sul mercato finestre di vantaggio competitivo sempre più brevi.

Il risultato di questo studio dimostra, sugli atteggiamenti dei diversi leader aziendali, come il loro sforzo per facilitare la connessione dei punti dei dati attraverso una varietà di asset e servizi, l'adozione di metodi di analisi dei dati e automatizzazione di processi tramite intelligenza artificiale, gli investimenti in multi-cloud, e la misurazione dell'impatto sul business dei loro investimenti in dati, ripaghino tutto lo sforzo fatto. Per cercare quindi di replicare il lavoro fatto da questi grandi leader, tutte le aziende e organizzazioni, dovranno prioritizzare l'imperativo dei dati di valore, se

vorranno rimanere sul panorama competitivo, dal momento che il cambiamento sarà sempre più veloce nell'era digitale.

1.4 Come estrarre valore dai dati

I dati, come già osservato in precedenza, rappresentano una importante fonte di valore aziendale. La vera sfida per le imprese sta nel riuscire ad estrarre questo valore dalle enormi quantità di dati che generano e condividono ogni giorno esternamente, ad esempio tramite lo scambio con altre fonti o clienti, oppure internamente, come i dati provenienti da differenti impianti e macchinari e che vengono poi raccolti ed elaborati ad esempio da un Data Warehouse aziendale. Uno dei modi principali per poter estrarre valore tramite i dati, è quello di fare analisi con essi. Queste analisi, anche conosciute in linguaggio internazionale come Analytics, possono essere di vario tipo. Esistono analisi real-time nelle operations aziendali, analisi predittive, analisi prescrittive, sentiment analysis, ecc... Dall'integrazione delle analisi real-time nelle operations nasce il concetto di operational analytics. L'operational analytics include attività e tecniche applicate a set di dati per poter analizzare le campagne di marketing, la penetrazione delle vendite, l'adozione dei prodotti, l'ottimizzazione delle risorse e la gestione dei rischi. Un processo di Operational Analytics prevede il monitoraggio e l'integrazione dei flussi di informazioni in tempo reale, la deduzione di conclusioni basate su modelli predittivi di comportamento e l'attivazione di risposte ed avvisi automatici. Una soluzione completa di operational analytics include la preparazione di dati storici per la precompilazione di modelli di comportamento. Questi calcoli predittivi, vengono generalmente eseguiti in anticipo in maniera tale da consentire risposte tempestive di fronte alla realizzazione degli eventi. Esempi di questo genere di modelli possono essere utilizzati in ambito retail oppure finanziario. Nel primo caso, un'analisi del comportamento dell'utente durante gli acquisti online, come valutazione dei prodotti nella lista dei preferiti, ricerche più frequenti oppure prodotti presenti nel carrello, può permettere di identificare i prodotti spesso acquistati congiuntamente. Per avere una previsione dell'andamento dei mercati finanziari, vengono valutate attentamente le informazioni di prezzo, tassi di interesse e tassi di variazione dei prezzi, ottenute tramite analisi dei dati storici. A questo punto, una volta sviluppati i modelli predittivi, in modo che siano utili e convenienti, vengono implementate soluzioni che integrano i dati storici e quelli generati attualmente, includendo anche dati real-time e in streaming, strutturati e non strutturati, in maniera tale da popolare i modelli predittivi e attivare azioni in base alle previsioni. La soluzione ottenuta dovrà garantire che i flussi di dati real-time che seguono le

regole del modello, vengano elaborate in maniera corretta e che vengano generate correttamente le risposte automatiche ad eventi significativi nei dati.

Lo sviluppo di modelli predittivi può essere chiamato Data Science, definita come tale poiché viene utilizzato il metodo scientifico per sviluppare e valutare il modello, infatti, all'interno di questa disciplina sono uniti data mining, analisi statistica e il machine learning con l'integrazione e modellazione di dati in maniera tale da costruire modelli predittivi che esplorino i pattern di contenuti dei dati. Con la Data Science viene sviluppata un'ipotesi che riguarda il comportamento osservabile nei dati prima di un determinato evento. Come già asservato, l'acquisto di un oggetto prevede, con una determinata probabilità, che ci sia in seguito l'acquisto di un altro oggetto, con caratteristiche più o meno simili. Ad esempio all'acquisto di una stampante segue generalmente l'acquisto di cartucce come ricarica per l'inchiostro. A questo punto viene analizzato un grosso quantitativo di dati storici per determinare quanto frequentemente le ipotesi si siano verificate nel passato, permettendo così di verificare l'accuratezza probabilistica del modello. Se una determinata ipotesi risulta validata con sufficiente frequenza e se il comportamento predetto risulta utile ai fini aziendali, allora il modello può diventare la base per un processo di intelligence operativa che possa permettere di predire i comportamenti futuri, eventualmente anche in tempo reale.

Addentrandosi più nel dettaglio, l'analisi predittiva è il sottocampo dell'apprendimento supervisionato in cui utenti cercano di modellare elementi di dati e prevedere i risultati futuri attraverso una valutazione delle stime di probabilità. L'analisi predittiva rappresenta lo sviluppo di modelli probabilistici basati su variabili, tra le quali sono inclusi anche i dati storici, relativi ad una grande varietà di possibili eventi (acquisti, variazioni di prezzo, variazione di volumi, ecc...). In particolare, quando il modello riceve altre informazioni attiva una reazione da parte dell'organizzazione. Il fattore di attivazione potrebbe essere un evento, ad esempio l'aggiunta di prodotti da parte di un cliente nel suo carrello durante lo shopping online. Altri esempi potrebbero essere l'arrivo di feed di notizie o dati provenienti da sensori, oppure ancora, una variazione della domanda di un determinato prodotto che necessiterebbe un repentino allineamento della produzione da parte degli impianti. Nel caso di società quotate in borsa, il fattore di attivazione esterno potrebbe avere notevole importanza. In questo caso particolare, le notizie che vengono riportate su una società sono un importante elemento premonitore del cambiamento nel prezzo delle sue azioni. Un'analisi predittiva sul movimento delle azioni dovrebbe quindi includere il

monitoraggio delle notizie, valutando se le notizie siano buone o cattive, permettendo così di prevedere un eventuale aumento o crollo del prezzo delle azioni.

Un fattore che va sicuramente monitorato all'interno di questi tipi di modello, è la risposta del sistema alla grande quantità di dati che arrivano in real-time. La quantità di tempo impiegata da parte di un modello predittivo tra la stima e la realizzazione dell'evento previsto è spesso molto piccola. Di fronte a determinate condizioni che lo necessitano, a volte l'investimento in soluzioni tecnologiche a bassissima latenza, che permettono quindi un trasferimento molto veloce di dati, come reti ad alta velocità oppure la vicinanza fisica all'origine dei dati, permette di ottimizzare la capacità da parte di un'organizzazione di reagire alla previsione.

L'analisi prescrittiva può essere considerata il proseguimento dell'analisi predittiva. Infatti, mentre l'analisi predittiva prevede soltanto i risultati delle azioni che si sono verificate, l'analisi prescrittiva definisce anche le azioni che potranno influire sui risultati. Ciò significa che questo tipo di analisi anticipa quello che accadrà, quando accadrà ed il possibile motivo per cui accadrà, provando ad individuare delle azioni in risposta ai risultati, risalendo alle possibili cause che si celano dietro ad un determinato evento. Dal momento che l'analisi prescrittiva può mostrare le implicazioni di varie decisioni, può suggerire come sfruttare una determinata opportunità o evitare certi rischi. Questa analisi può continuamente raccogliere dati per poter predire e prescrivere nuovamente, in modo tale da migliorare la precisione della previsione e risultare in prescrizioni migliori.

Nel corso del tempo sono stati svolti numerosi studi per cercare di dimostrare un effettivo valore aziendale portato dai dati ed in particolar modo dalle attività di analisi di dati e Big Data. In un articolo di Harvard Business Review, Brynjolfsson e McAfee hanno notato come le aziende "data driven", ovvero aziende che impostano molto del loro business basandosi sulla raccolta ed elaborazione di dati, sono il 6% più redditizie e il 5% più produttive rispetto ai loro concorrenti. Inoltre, un'efficace combinazione di infrastrutture, risorse umane e capacità di gestione delle aziende è significativa per l'adozione di successo di analytics e prestazioni finanziarie/operative superiori. La sfida tecnica dell'utilizzo di dati è sicuramente reale, ma la sfida manageriale è ancora più grande, affrontando dall'alto verso il basso la gerarchia organizzativa. Per far fronte a questa sfida, Brynjolfsson e McAfee suggeriscono inoltre cinque aree per le aziende su cui concentrarsi maggiormente, ovvero tecnologia, leadership, processo decisionale, gestione dei talenti e cultura aziendale. Naturalmente le aziende non possiedono sempre tutte le capacità necessarie per

affrontare queste sfide, quali le persone, gli strumenti, la focalizzazione sull'intenzione dei dati e le informazioni analitiche giuste.

Uno studio interessante di Amankwah-Amoah and Adomako propone una relazione tra il possibile fallimento aziendale e l'utilizzo di analisi di dati e big data da parte delle imprese. In generale, il fallimento aziendale può essere ricondotto a molte cause differenti, una tra queste potrebbe essere per esempio, l'accumulo di decisioni sbagliate ed errori strategici. Il fallimento aziendale può essere visto attraverso due prospettive: punti di vista endogeni ed esogeni. La prospettiva endogena esamina le azioni fatte e quelle non fatte dei manager, la loro abilità e lo stile di leadership come cause di fallimento aziendale. Una linea di ricerca ha anche prestato attenzione alla mancanza di dipendenti altamente qualificati e di strategie aziendali come componenti dei fattori interni. Queste carenze spesso portano a una mancanza di creatività e unità di business intelligence limitate o inesistenti, creando così le condizioni per il fallimento aziendale.

Il fallimento aziendale può anche derivare da conflitti intra-organizzativi che deviano le risorse e l'attenzione verso la risoluzione dei conflitti, creando così la possibilità che gli errori prosperino. La prospettiva esogena si basa sul presupposto che il fallimento derivi da fattori esterni incontrollabili come il cambiamento tecnologico, le azioni delle imprese rivali, la concorrenza, la fortuna e lo stato dell'economia. Gli studi, nel corso degli anni, hanno dimostrato come una nuova tecnologia ed un possibile cambiamento di paradigma tecnologico, possano rendere obsolete molte fonti tradizionali di vantaggio competitivo. Questo problema è importante dato che la tecnologia e l'introduzione di nuove tecnologie potrebbero esercitare pressioni diverse sulle imprese, portando alla chiusura dell'attività. Negli ultimi anni, il consolidamento globale delle industrie ha scatenato nuove fonti di concorrenza, esercitando pressioni sulle imprese e portando al loro collasso. L'intensa concorrenza in un mercato popolato da una grandissima quantità di industrie significa meno spazio per errori strategici ed errori che portano ad un'allocazione errata di risorse limitate. Poiché la concorrenza si intensifica in una serie di settori, è diventato un imperativo strategico per le aziende che cercano un vantaggio competitivo, stringere relazioni più strette con i principali stakeholder e rimanere in sintonia con le richieste dei clienti.

Storicamente, l'accesso ai dati dei clienti e ai record delle aziende era limitato dai costi e dallo spazio di archiviazione estremamente elevati. L'accumulo di dati che potrebbero equipaggiare le aziende per la concorrenza di mercato tendeva a essere un processo lungo e costoso. Dall'inizio di questo secolo, il forte calo dei costi di archiviazione dei dati, insieme ai progressi tecnologici, ha inaugurato

una nuova era in cui la capacità delle aziende di raccogliere dati voluminosi e archivarli è aumentata enormemente.

Una delle sfide principali per le aziende è quella di essere abili nel trasformare i dati in informazioni per generare intuizioni nella risoluzione dei problemi organizzativi. Analisi recenti hanno dimostrato come il semplice possesso di dati ad alta varietà e velocità non produca di per sé un vantaggio competitivo sostenibile, piuttosto è la capacità di assemblare questi dati strutturati e non strutturati, analizzarli e utilizzare le intuizioni per informare le decisioni delle organizzazioni che producono tale vantaggio. Di conseguenza, fare affidamento esclusivamente sulla raccolta di grandi volumi di dati, senza passare all'analisi per informare le decisioni organizzative è più probabile che porti a un sottoutilizzo delle risorse e di conseguenza al fallimento aziendale. La capacità di analisi dei dati è quindi concettualizzata, per riferirsi in generale alla capacità delle aziende di trarne vantaggio per sviluppare capabilities che portino ad un vantaggio competitivo sostenibile. Il nocciolo della questione è che, sviluppando capacità di analisi dei dati, le aziende sono maggiormente in grado di trarre pieno vantaggio da altre competenze funzionali, comprese le capacità di marketing e la consegna del prodotto. Per le aziende in cerca di risorse, i dati sono una delle risorse sottoutilizzate che può essere sfruttata per avere accesso a nuove conoscenze e migliorare quindi la qualità delle decisioni. Sfruttando i dati, le aziende sarebbero in grado di mettere alla luce conoscenze nascoste, scoprire modelli prima dei concorrenti e, infine, prendere decisioni più affidabili in maniera più veloce.

Diventando sempre più proattive nello sfruttare i dati, le organizzazioni migliorano la loro prontezza strategica e sono in grado di adottare una strategia più efficace sulla scia dei cambiamenti ambientali che minacciano o indeboliscono le competenze aziendali. Mobilitando e analizzando i dati, le aziende sono in grado di migliorare l'esperienza dei clienti e la loro efficienza complessiva. Il fallimento aziendale deriva dalla capacità o incapacità di mobilitare e utilizzare efficacemente le risorse e le capacità dell'azienda per rispondere alle minacce derivanti dall'ambiente esterno. In quanto tale, il vantaggio competitivo in molti settori si è spostato verso lo sviluppo ed il possesso di capitale umano, che è alla base delle prestazioni delle imprese.

Per valutare quindi, come si possa estrarre il valore dai dati, è importante comprendere che il semplice possesso o l'accesso ai big data non produca automaticamente successo o vantaggio competitivo, piuttosto è la capacità di estrarre e utilizzare i dati che può fornire alle aziende migliori strumenti per mitigare il fallimento aziendale e migliorare la loro competitività. Questo risultato è

coerente con argomentazioni teoriche consolidate nella strategia secondo cui è improbabile che il semplice possesso di un'attività produca un vantaggio competitivo, ma piuttosto, che questo possa nascere dalla capacità di utilizzare l'attività stessa.

Da quello detto fino ad ora si possono cogliere alcune implicazioni pratiche. Innanzitutto, l'era dei big data richiede agilità strategica, in cui le aziende scansionano continuamente l'ambiente aziendale per aggiornare le proprie conoscenze e competenze, e fare in modo che la propria strategia sia allineata e coerente con l'utilizzo di nuovi set di dati e informazioni. Sebbene raggiungere tale agilità strategica possa essere costoso, rappresenta un utilizzo efficiente delle risorse a lungo termine, dato il risparmio derivante da errori strategici, allocazione non efficiente delle risorse e bassa attenzione organizzativa.

2 IL DATA MANAGEMENT

2.1 Definizione di Data Management

Nel corso del tempo il Data Management è stato oggetto di diverse definizioni. Tra queste risulta piuttosto rilevante quella fornita da DAMA international, una delle più importanti organizzazioni nel panorama internazionale per il Data Management. DAMA sostiene che il Data Management sia “lo sviluppo, l’esecuzione, e la supervisione di piani, politiche, programmi e procedure che garantiscono, controllano, proteggono e migliorano il valore dei data asset e informazioni per tutta la durata del loro ciclo di vita”. Come si può dedurre da questa definizione, le attività di Data Management sono di ampia portata. Comprendono aspetti anche molto diversi tra di loro, partendo dalla possibilità di prendere decisioni coerenti su come ottenere valore strategico dai dati fino ad aspetti più tecnici di implementazione e prestazione dei database o dell’architettura di sistema. Un concetto fondamentale che si evince subito da questa definizione, è che l’applicazione delle pratiche di Data Management in una organizzazione coinvolgono sia comparti aziendali di business sia presidi di information technology, facendo in modo che le persone in entrambe le aree collaborino per garantire la disponibilità di dati di qualità adeguata in grado di rispondere alle esigenze strategiche, tattiche e operative. Sebbene le aree aziendali siano diverse tra di loro, coinvolgendo team diversi, con skills e background differenti, i dati gestiti in una singola area non sono completamente disgiunti da quelli delle altre. Riconoscendo le sinergie tra informazioni provenienti da diversi comparti, si genera tra le aree aziendali una connessione proprio attraverso i dati. In effetti, l’obiettivo finale del Data Management è consentire alle organizzazioni di estrarre valore dai propri data asset, proprio come un management efficace delle risorse finanziarie e fisiche consente alle organizzazioni di ottenere valore da altri asset aziendali.

Uno dei primi requisiti per raggiungere questo obiettivo è la conoscenza dei dati di cui dispone un’organizzazione. Come tanti altri processi di management è necessario bilanciare le esigenze strategiche e operative, risultati attesi e risorse necessarie per ottenerli. La determinazione e l’attuazione di questo equilibrio è sostenuta da una serie di principi, prospettive, pratiche di Data Management e ne guidano la sua pratica. I principi riconosciuti dal DAMA sono i seguenti:

- I dati sono un asset con proprietà uniche: i dati sono un asset, ma si differenziano dagli altri asset per i modi che riguardano le modalità di gestione. Per chiarire meglio questo concetto

analizziamo la differenza tra un asset fisico come il petrolio e un asset come i dati. Partendo dalla fornitura, la quantità di petrolio disponibile sul nostro pianeta è limitata mentre i dati sono virtualmente infiniti. Il petrolio è una risorsa limitata anche per il fatto che non può essere copiato o replicato, mentre i dati possono esserlo. Il controllo del petrolio è nelle mani di pochi attori principali a livello mondiale, invece nessuno sa chi controlli o detenga i nostri dati personali. Una volta utilizzato, il petrolio non può più essere utilizzato nuovamente, i dati invece possono essere usati senza limitazioni e senza paura che possano consumarsi.

- Il valore dei dati può e deve essere espresso in termini economici: definire i dati come asset implica che essi abbiano un valore. Quest'ultimo però, sebbene esistano tecniche per misurare il valore qualitativo e quantitativo dei dati, non ha degli standard per farlo in maniera accurata. Uno dei modi dai cui un'organizzazione può partire con la misurazione del valore dei dati, è la valutazione dei costi che possono essere imputati alla bassa qualità dei dati e nel contempo, i benefici derivanti da quelli di alta qualità.
- Gestire i dati significa gestire la qualità dei dati: per gestire la qualità dei dati è fondamentale che le organizzazioni comprendano a fondo i requisiti degli stakeholder in merito alla qualità e misurare i dati a fronte di questi requisiti.
- Per gestire i dati occorrono metadati: i dati che vengono utilizzati per gestire ed utilizzare i dati stessi sono definiti metadati. Questi supportano svariati processi legati alla creazione, elaborazione ed utilizzo dei dati, come ad esempio architettura, modeling, stewardship, governance, data quality management, system development, business operations e IT analytics.
- La gestione dei dati richiede pianificazione: i dati nascono in diversi punti e si trasferiscono in altri punti per l'utilizzo. Per tenere sotto controllo gli step evolutivi del dato e gli attori coinvolti, è necessaria una pianificazione sia architetturale che processuale.
- Il data management è inter-funzionale, richiede un ventaglio di abilità e competenze: dal momento che un singolo team non può gestire tutti i dati di un'organizzazione, il Data Management richiede competenze tecniche e non tecniche unite dalla capacità di collaborazione.
- Il Data Management richiede una prospettiva aziendale: il Data Management deve essere applicato a tutta l'azienda per essere il più efficace possibile. Da qui nasce il bisogno di Governance, che risulta quindi intrecciata con la gestione dei dati.

- Il Data Management deve tenere conto di un ventaglio di dinamiche: questa disciplina deve evolversi continuamente per stare al passo con i modi con cui i dati vengono creati e utilizzati, e al tempo stesso con coloro che utilizzano i dati stessi.
- Il Data Management implica la gestione del ciclo di vita dei dati: i dati seguono percorsi e attraversano fasi che variano nel tempo; la gestione dei dati richiede la gestione di questo ciclo di vita, considerando anche la complessità che ne consegue, come il fatto che i dati generino altri dati e che i dati vengano modificati durante la loro vita.
- Tipi diversi di dati hanno caratteristiche di ciclo di vita diverse: le pratiche di Data Management devono essere abbastanza flessibili da soddisfare diversi requisiti legati al ciclo di vita in funzione del tipo di dato.
- Gestire i dati implica gestire i rischi associati ad essi: i dati possono infatti essere persi, sottratti o utilizzati impropriamente. I rischi derivanti dall'uso dei dati o da eventi che li riguardano devono essere considerati come parte integrante della loro gestione.
- I requisiti di Data Management devono pilotare le scelte dell'Information Technology: una gestione sapiente del dato richiede un approccio che garantisca che la tecnologia sia al servizio e non che determini l'esigenza di dati strategici dell'organizzazione.
- Un Data Management efficace richiede leadership commitment: dal momento che il Data Management coinvolge un insieme complesso di processi, che richiedono coordinamento, collaborazione e commitment, sono necessarie non solo abilità manageriali, ma anche visione e scopo che accompagnino una committed leadership.

2.2 Knowledge areas

Per poter distinguere e lavorare al meglio categorizzando le varie aree su cui il Data Management deve operare all'interno di un'organizzazione, il DAMA propone un framework nel quale vengono approfondite le knowledge areas che costituiscono l'ambito generale del Data Management. Una raffigurazione comune nel settore è rappresentata dalla DAMA Wheel che definisce appunto le knowledge areas del Data Management. In questo tipo di rappresentazione si può notare come la Data Governance venga posta al centro delle attività di Data Management, dal momento che è fondamentale la presenza della governance per garantire coerenza interna ed equilibrio tra le varie funzioni. Le altre knowledge areas sono disposte attorno alla governance, definendo così la caratteristica "Wheel". Sono tutte discipline e pratiche necessarie per poter esercitare una Data Management matura, ciascuna applicabile in momenti diversi, in base ai requisiti dell'organizzazione.

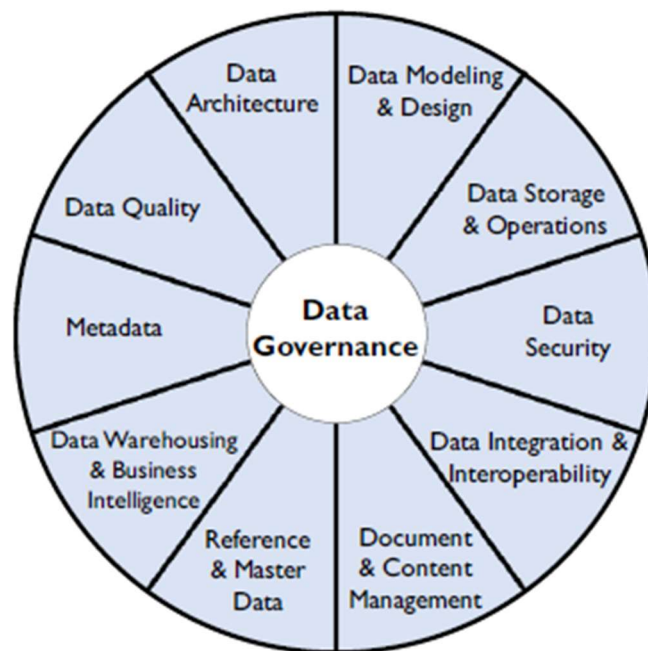


Figura 1 DAMA Wheel. Fonte: Dmbok data management body of knowledge 2nd edition (2020)

2.2.1 Data Governance

Una tra le definizioni più significative della Data Governance considera quest'ultima come l'esercizio dell'autorità e del controllo (pianificazione, monitoraggio e applicazione) nella gestione dei data asset. La funzione di Data Governance risulta fondamentale in quanto guida tutte le altre funzioni di Data Management. Il suo scopo è quello di assicurare che i dati siano gestiti adeguatamente, in accordo con policy e best practices (Ladley, 2012). La Data Governance rientra nel contesto generale del Data Management, il cui obiettivo è invece, come già visto, assicurare che l'organizzazione ottenga valore dai propri dati. La Data Governance si concentra quindi su come persone e processi si comportano in relazione ai dati. Le caratteristiche di un programma di Data Governance non sono definite a priori ma dipendono strettamente dalle necessità dell'organizzazione. Per raggiungere questi obiettivi la Data Governance svilupperà policy e procedure, incoraggerà pratiche di data stewardship all'interno dell'organizzazione nelle varie aree e ai diversi livelli. Il processo di applicazione della Data Governance sarà orientato verso un cambiamento che renda concretamente percepibile all'organizzazione i benefici di una gestione ordinata dei dati ed i comportamenti che ne conseguono. Le organizzazioni, proprio per raggiungere questo obiettivo, spingono in maniera

sempre crescente per diventare data-driven. Definiscono le loro strategie considerando in modo consapevole i dati aziendali come un fattore abilitante imprescindibile per la loro realizzazione.

Un aspetto però da tenere sicuramente in considerazione è il fatto che questa visione impone un cambiamento di paradigma che coinvolge tutti i livelli aziendali. Valutando che una strategia è implementabile solo se esistono le condizioni culturali per metterla in opera, l'insediamento di processi, ruoli, tecniche, strumenti di Data Governance deve necessariamente prevedere interventi finalizzati alla gestione di un cambiamento culturale, supportati da un adeguato leadership commitment. Per molte organizzazioni il cambiamento culturale risulta la sfida maggiore. Uno dei principi fondamentali, secondo Hiatt e Creasey, nella gestione del cambiamento, è che il cambiamento organizzativo richiede un cambiamento individuale. Quindi, se i processi di Data Management introducono cambiamenti operativi importanti, la Data Governance deve regolarne l'articolazione in modo da strutturare per quanto possibile i necessari cambiamenti, ma deve anche sostenere con opportune azioni di formazione, informazione, addestramento la transizione verso il nuovo assetto dei processi stessi.

Uno dei principi fondamentali è che la Data Governance deve essere necessariamente allineata alla strategia aziendale. Più chiaramente aiuta a risolvere i problemi organizzativi e più persone cambieranno il proprio comportamento per adottare le pratiche di governance. Tra i driver principali del governo del dato vi sono la riduzione del rischio e il miglioramento, in termini di performance e efficienza, dei processi. La Data Governance è un presidio collettivo dei data asset, non può coinvolgere solo una struttura dedicata. Piuttosto si esprime come un insieme di comportamenti e pratiche condivisi da tutti i comparti aziendali. L'Information Technology gioca sicuramente un ruolo molto importante ma bisogna sempre tener presente che il processo è business-driven, non IT-driven. Questo significa che l'Information Technology concorre a sostenere pratiche e modelli di Data Governance sulla base di logiche dettate dai requisiti di business.

Per comprendere meglio questi concetti, si può equiparare la Data Governance all'auditing e alla contabilità. Come un auditor supervisiona e controlla i processi finanziari, ma non esegue fisicamente la gestione finanziaria, allo stesso modo la Data Governance assicura che i dati siano gestiti in maniera corretta, senza dover intervenire direttamente nella gestione del dato, che risulta invece materia di applicazione del Data Management stesso. Come mostrato in figura, si può quindi notare come la Data Governance rappresenti una "intrinseca separazione delle responsabilità tra controllo ed esecuzione".

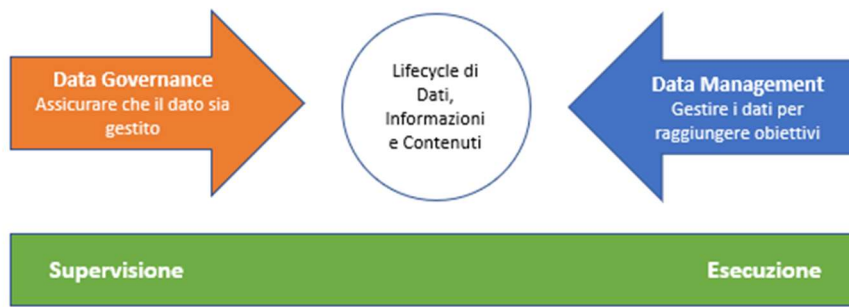


Figura 2 Data Governance e Data Management. Fonte: Dmbok data management body of knowledge 2nd edition (2020)

Per definire la responsabilità per dati e processi che assicurino un effettivo controllo ed utilizzo degli data asset, si può utilizzare il termine Data Stewardship. Lo scopo delle attività di stewardship sarà differente da un'organizzazione all'altra in base a strategie, organizzazione, cultura, problemi che l'organizzazione stia tentando di risolvere, livello di maturità nella gestione e altri fattori. In generale, viene definito "steward" una persona il cui compito consiste nel gestire le proprietà di un'altra persona. Un Data Steward analogamente gestisce invece i data asset per conto e nell'interesse dell'organizzazione. Queste figure inoltre sono responsabili per le attività di Data Governance, rappresentano gli interessi di tutti gli stakeholders ed è necessario che considerino un punto di vista globale per assicurare che i dati dell'azienda siano di alta qualità ed usati in maniera efficace. Le figure principali riconosciute come steward sono ad esempio Chief Data Steward (CDO), Executive Data Steward, Enterprise Data Steward, Business Data Steward, Data Owner, Technical Data Steward, Coordinating Data Steward.

2.2.2 Data Architecture

In generale con architettura si intende una disposizione organizzata di elementi destinati a ottimizzare funzione, performance, fattibilità, costi ed estetica di una struttura complessiva o di un sistema. Le pratiche che riguardano questa disciplina, se ben gestite, permettono alle organizzazioni di comprendere lo stato attuale dei propri sistemi, in modo tale da incentivare il cambiamento desiderabile verso lo stato futuro, consentire conformità normativa e migliorare l'efficacia. Una gestione efficace dei dati e dei sistemi in cui sono archiviati e utilizzati è un obiettivo comune delle pratiche legate alla Data Architecture. La Data Architecture assume il massimo valore quando supporta completamente le esigenze dell'intera azienda, in particolare un suo dominio: l'Enterprise Data Architecture consente una costante standardizzazione e integrazione di dati in tutta l'azienda.

I Data Architect creano e gestiscono le conoscenze organizzative dei dati e dei sistemi attraverso i quali essi si muovono, identificando le opportunità di utilizzo degli stessi dati e permettendo così di ottenere una riduzione dei costi ed una mitigazione dei rischi. Tendenzialmente nello sviluppo di un'architettura aziendale non si parte mai da zero, ma si cerca sempre di analizzare gli elementi già esistenti permettendo così di implementare il modello in maniera incrementale, iniziando quindi con il modello aziendale esistente e cercando di migliorarlo in maniera continua. In ogni caso sarebbe sempre meglio sviluppare da subito la Data Architecture in modo da poter poi anticipare i possibili problemi ed errori.

La Data Architecture opera ed è composta da quattro differenti domini, tra di loro connessi, influenzandosi e vincolandosi a vicenda. Sono business, dati, applicazioni e tecnologia. Enterprise Business Architecture identifica in che modo un'azienda crea valore per i clienti e gli altri stakeholder, utilizzando come elementi modelli di business, processi, servizi, strategie. Stabilisce inoltre i requisiti per gli altri domini di architettura. Enterprise Data Architecture descrive come devono essere organizzati e gestiti i dati. Gli elementi coinvolti sono per esempio Data model, definizioni dei dati, flussi di dati, che permettono di gestire i dati creati e richiesti dall'architettura aziendale, realizzati attraverso figure di data architect, data modeler e data steward. L'architettura delle applicazioni aziendali descrive la struttura e la funzionalità delle applicazioni in un'azienda e agisce su dati specifici in base alle esigenze del business, utilizzando sistemi di business, pacchetti software e database. Infine l'architettura della tecnologia aziendale descrive la tecnologia fisica necessaria per consentire ai sistemi di funzionare nel modo corretto, attraverso piattaforme tecniche, reti, sicurezza e strumenti di integrazione.

La progettazione di una Enterprise Data Architecture include la rappresentazione dei dati di business, quindi operazioni di raccolta, archiviazione, integrazione, trasferimento e distribuzione dei dati stessi. Quando i dati si muovono all'interno di un'organizzazione, attraversano una serie di fasi prima di essere consegnati agli stakeholder, tra cui protezione, integrazione, archiviazione, condivisione, catalogazione, registrazione, analisi. Lungo il percorso i dati possono ricevere verifiche, miglioramenti, essere collegati e aggregati. Per questo motivo la descrizione dell'Enterprise Data Architecture deve essere completa e comprensibile e deve includere modelli di Enterprise Data, ad esempio strutture e specifiche di dati, entità dati aziendali, le loro relazioni, le regole di business di guida e attributi. Deve essere inclusa nell'architettura anche la progettazione

del flusso di dati, che mappa il trasferimento dei dati attraverso processi e posizioni permettendo ad esempio di analizzare il processing su database, applicazioni, piattaforme e reti.

2.2.3 Data Modeling and Design

Quando si parla di Data Modeling, tendenzialmente si intende il processo di scoperta, analisi e studio dei requisiti dei dati e la loro successiva rappresentazione in un determinato formato definito appunto Data Model. Quest'ultimo risulta fondamentale all'organizzazione per la comprensione dei propri data asset. Il Data Modeling rappresenta un tassello fondamentale nella gestione dei dati, anche per il fatto che il processo di modeling supporta le organizzazioni nella determinazione e nella rappresentazione dei dati e dei relativi attributi, nonché del modo in cui i dati sono in relazione l'uno con l'altro. Un elemento fondamentale per gli utilizzatori dei dati stessi sono i metadati, che sono parte integrante di un data model. Questi sono fondamentali per altre funzioni di Data Management, come ad esempio, la definizione del data lineage, necessario a loro volta per l'analisi dei dati stessi.

La maggior parte dei data model contiene gli stessi elementi costitutivi di base: entità, relazioni, attributi e domini. Le entità possono essere considerate come una risposta a una domanda fondamentale (chi, cosa, quando, dove, perchè o come), oppure ad una combinazione di queste domande. Alcuni esempi di entità sono dipendente, paziente, cliente per la domanda "chi?". Prodotto, servizio, materia prima per rispondere alla domanda "cosa?". Ora, data, periodo fiscale per rispondere alla domanda "quando?". Indirizzo, URL del sito Web, punto di distribuzione per rispondere alla domanda "dove?". Le entità sono collegate tra di loro tramite associazioni definite relazioni. Per misurare, identificare o descrivere un'entità vengono utilizzate proprietà definite attributi. Un prodotto finito porta con se quindi attributi come ad esempio nome, dimensioni, peso, data di produzione, ora di produzione. Infine il dominio è definito come l'insieme completo di possibili valori assegnabili a un attributo. Considerando per esempio l'attributo ora di produzione riferito ad un determinato prodotto, questo avrà un dominio, come la fascia oraria di produzione di quel determinato pezzo. I valori all'interno del dominio sono considerati validi, mentre quelli al di fuori del dominio, non validi.

I motivi per cui i Data Model sono così importanti per una gestione efficiente dei dati sono molteplici. Questi forniscono un vocabolario comune inerente ai dati, acquisiscono e documentano la conoscenza esplicita dei dati e dei sistemi di un'organizzazione, sono lo strumento principale di

comunicazione durante i progetti. In un Data Model esistono diversi livelli di astrazione e tipi di organizzazione dei dati. Quelli utilizzati più comunemente sono: relazionale (relational), dimensionale (dimensional), orientato agli oggetti (object-oriented), basato sui fatti (fact-based), basato sul tempo (time based) e noSQL. Esistono tipologie di data model enterprise, conceptual, logical e physical. A livello concettuale il data model contiene le entità e le relazioni tra le entità, per esempio un macchinario (entità1) produce (relazione) un determinato prodotto (entità2). Il livello logico rappresenta una estensione del modello concettuale attraverso l'aggiunta di attributi, per esempio nome del macchinario di produzione, nome del prodotto e codice del prodotto. Il livello fisico rappresenta una soluzione tecnica dettagliata, adattando un data model logico per operare all'interno di un set di strumenti hardware, software e di rete.

L'obiettivo del Data Modeling è quello di confermare e documentare la comprensione di diversi punti di vista circa i dati gestiti in un certo ambito (una applicazione, un processo, un'area aziendale) facilitando la realizzazione di applicazioni IT coerenti con la visione e le esigenze di business, e creando così le basi su cui costruire iniziative di ampia portata come ad esempio programmi di Data Governance o di Master Data Management. L'applicazione di pratiche e tecniche di Data Modeling ha impatti positivi sulla riduzione dei costi di implementazione e manutenzione, sull'efficacia del supporto ai processi aziendali, sulla riusabilità di dati e dei relativi metadati nel futuro per poter così ridurre costi di sviluppo di nuove applicazioni.

Il contesto principale in cui viene eseguito più frequentemente il Data Modeling è quello del ciclo di vita dello sviluppo del sistema SDLC (System Development Life Cycle), ma può anche essere eseguito in progetti di portata più ampia in cui il risultato è una comprensione accurata dei dati organizzativi. In generale, il modello può includere al suo interno uno o più diagrammi, che fanno a loro volta uso di simboli standard che permettono di comprenderne in modo chiaro il contenuto. Questi simboli cercano di rappresentare i requisiti dei dati così come comunicati dal data modeler. I Data Model sono quindi il mezzo principale utilizzato per comunicare i requisiti dei dati tra business e l'IT. Ad esempio un Data Steward potrebbe essere Data Modeler, ponendosi in comunicazione tra queste due diverse aree aziendali.

2.2.4 Data Storage and Operations

Data Storage e Operations riguarda la progettazione, l'implementazione e il supporto dell'archiviazione dei dati. In generale un database può essere definito come una raccolta di dati memorizzati, a prescindere dalla struttura o dal contenuto. In questa knowledge area del Data Management giocano un ruolo fondamentale gli amministratori di database (DBA), che gestiscono operativamente gli interventi tecnici sulle basi dati (backup, restore, partizionamento, ...). Le attività di Data Storage e Operations sono fondamentali per le organizzazioni che fanno affidamento sui dati. I loro principi cardine riguardano la gestione della disponibilità dei dati per tutto il ciclo di vita, la garanzia d'integrità dei data asset e la gestione delle prestazioni che riguardano le transazioni di dati.

Dal punto di vista della organizzazione fisica, possiamo distinguere due tipologie principali di database: centralizzato e distribuito. I database centralizzati mantengono tutti i dati in un unico sistema e in un unico luogo; quindi per accedere ai dati tutti gli utenti si rivolgeranno a quell'unico sistema. La centralizzazione potrebbe comportare dei rischi per i dati che devono essere ampiamente disponibili, infatti, se il sistema centralizzato non fosse più disponibile, non esisterebbero altre alternative per accedere ai dati. I database distribuiti si basano invece sulla distribuzione degli storage in molti nodi locali. Esiste poi una terza tipologia di database definiti federati. In questa configurazione i database, spesso separati geograficamente, sono connessi tra loro tramite una rete, rimangono autonomi ma partecipano ad una federazione in modo tale da permettere una condivisione parziale e controllata dei propri dati. Grazie alla federazione non si ha una reale integrazione di dati nei database costituenti, ma l'interoperabilità dei dati consente di vedere i database federati come un unico elemento.

Esistono due modelli fondamentali di database processing: ACID e BASE. Questi rappresentano due estremi opposti in quanto ACID "Atomicità, Coerenza, Isolamento, Durabilità" è una tipologia che viene utilizzata da decenni. Questo modello ha come linguaggio di manipolazione di riferimento l'SQL (Structured Query Language) e garantisce che una transazione, ovvero l'esecuzione di una serie di operazioni sui dati che devono essere eseguite contemporaneamente, venga svolta in modo consistente, per evitare incoerenze tra dati (ad esempio la presenza di una riga di un'ordine di acquisto senza che esista una testata dell'ordine stesso). I sistemi di tipo BASE sono invece diffusi comunemente negli ambienti dei Big Data: la diffusione senza precedenti di volumi e variabilità dei dati, il bisogno di archiviare dati non strutturati ed il crescente bisogno di flessibilità per quanto

riguarda scalabilità, progettazione, processing, costo e disaster recovery ha dato vita al concetto opposto di ACID, ovvero BASE “Basically Available, Soft State, Eventual consistency”. In questo processo il sistema continua a ricevere input ma non verifica la coerenza di ogni transazione prima di passare a quella successiva. Controllare la coerenza di ogni singola transazione in ogni momento di ogni modifica, aggiunge costi giganteschi a un sistema che processa milioni di transazioni ogni momento. Per questo motivo le grandi organizzazioni online e le società di social media utilizzano comunemente questo processing, dal momento che la precisione immediata di tutti gli elementi di dati in qualsiasi momento non è necessaria.

2.2.5 Document and content management

Quando si parla di dati occorre fare una distinzione tra dati strutturati e dati non strutturati. Nel primo caso si considerano i dati che sono organizzati secondo schemi definiti, e che per questo agevolano la gestione delle relazioni tra informazioni. I dati non strutturati sono invece tutti quei dati che non seguono uno schema preciso, una struttura che li suddivide in campi con significati prestabiliti: sono dati non strutturati file multimediali quali foto e video, oppure file di testo. Come si può vedere dal grafico, nel corso degli anni si è potuta osservare una crescita sempre maggiore del volume di dati non strutturati. Questo sta creando una forte sfida per le organizzazioni nella gestione dei dati non strutturati.

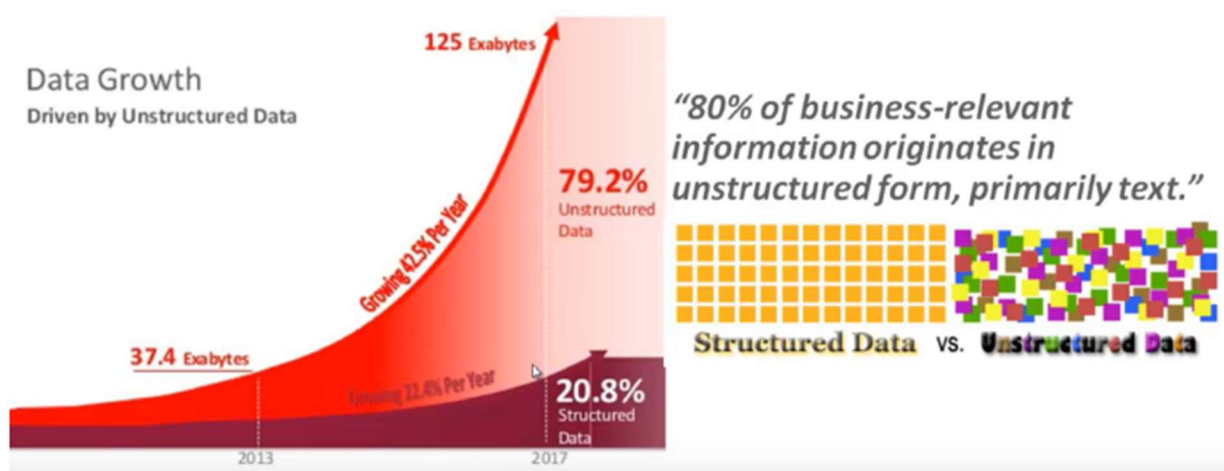


Figura 3 Charlie Berger - <https://blogs.oracle.com/>

La knowledge area Document and Content Management riguarda il controllo dell’acquisizione, mantenimento, accesso ed utilizzo dei dati e delle informazioni che coinvolgono dati non strutturati.

Il compito principale è quello di preservare l'integrità dei documenti e di altre informazioni non strutturate o semi strutturate e nel contempo permetterne l'accesso. Questo presidio potrebbe avere finalità anche strategiche, in molte organizzazioni i dati non strutturati hanno una relazione diretta con quelli strutturati. Tra i compiti principali del Document and Content Management possiamo citare la ricerca e l'uso efficaci ed efficienti di informazioni in formati non strutturati, l'integrazione tra dati non strutturati e dati strutturati, la definizione di policy aziendali interne per la gestione e prioritizzazione dei dati non strutturati.

Un esempio di disciplina importante in questo ambito di applicazione risulta quella di e-discovery. Con questo termine si intende il processo di ricerca di informazioni all'interno grandi volumi di dati non strutturati archiviati elettronicamente, che potrebbero essere utilizzati per esempio come prova in caso di azione legale, oppure a supporto di analisi di mercato. Lo sviluppo delle tecnologie dell'informazione ha determinato una crescita esponenziale del volume di informazioni state archiviate elettronicamente. La capacità da parte di un'organizzazione di rispondere a richieste di e-discovery, dipende da quanto abbia gestito efficacemente informazioni quali mail, chat, pagine web e in generale documenti elettronici, oltre all'utilizzo sapiente di dati e metadati.

Una maggiore efficienza è un aspetto fondamentale per il miglioramento del Document and Content Management. I progressi tecnologici nella gestione dei documenti stanno permettendo alle organizzazioni di linearizzare i processi, gestire i workflow, eliminare compiti manuali ripetitivi e favorire la collaborazione. Questi tipi di tecnologie permettono inoltre alle persone di individuare, accedere e condividere i documenti in maniera più veloce, oltre a prevenirne la perdita. Aspetto infine molto importante risiede nel fatto che vi sia anche un risparmio economico derivante da un utilizzo minore dello spazio occupato da archivi fisici e dalla riduzione dei costi per la gestione dei documenti.

2.2.6 Data Security

La disciplina di Data Security include la pianificazione, lo sviluppo e l'esecuzione di policy e procedure di sicurezza volte a garantire una corretta autenticazione, autorizzazione, accesso e verifica degli data asset. Proprio da qui nasce lo scopo delle pratiche di Data Security, ovvero proteggere i data asset e le informazioni, rispettando le leggi sulla privacy e riservatezza, le clausole contrattuali e le esigenze di business. Le esigenze di security possono nascere per diverse ragioni: dai bisogni degli stakeholder in termini di privacy e riservatezza, dalle leggi che esistono a tutela

degli interessi delle persone, da aspetti di business che coinvolgono il vantaggio competitivo stesso dell'impresa. Se infatti alcuni dati venissero rubati o distrutti dall'organizzazione, questo potrebbe rappresentare un danno di notevole importanza. Inoltre potrebbero esserci obblighi contrattuali relativi alla sicurezza stessa dei dati. I rischi coinvolgono gli aspetti sopra citati e la violazione di conformità legislativa o di obblighi contrattuali può comportare multe e sanzioni. Le violazioni ai dati possono causare perdita di reputazione per l'impresa e di fiducia da parte dei clienti. Il primo step per garantire la sicurezza delle informazioni è quello di classificare i dati dell'organizzazione in modo da identificare i dati che hanno bisogno di essere protetti. Dopo aver identificato e classificato i dati critici, questi vanno localizzati all'interno dell'organizzazione. Vanno determinati i rischi connessi alle possibili minacce, dividendole tra interne (dati che vengono persi o danneggiati da dipendenti o processi) ed esterne (hacker, organizzazioni criminali che cercano volontariamente di accedere ai dati o di corromperli). Vanno identificate e messe in opera le azioni a mitigazione di questi rischi.

I requisiti di sicurezza e sensibilità stabiliti da norme e regolamenti interni o esterni possono essere identificati anche a livello di insiemi omogenei di dati e non di singolo dato. Un approccio vincente per la gestione dei dati sensibili è l'utilizzo di metadati. A questo scopo risultano molto utili tecnologie esistenti che permettono di taggare i dati con metadati che li classifichino rispetto ai criteri di sicurezza. Attraverso lo sviluppo di un archivio centralizzato, tutte le aree aziendali possono sapere con precisione il livello di protezione che una determinata informazione sensibile debba avere. In questo caso l'applicazione di uno standard comune permette ai vari dipartimenti e business unit di usare gli stessi metadati e l'utilizzo di metadati di sicurezza standard. Consente di ottimizzare la gestione della protezione. Infine una guida all'utilizzo dei dati e processi di supporto tecnico, contribuisce all'abbassamento dei costi, aiutando a prevenire accessi non autorizzati e l'abuso dei dati. Dal momento che, con un dato sensibile le organizzazioni possono creare un rapporto di fiducia con clienti e partner, i metadati relativi alla sicurezza possono diventare un asset strategico per incrementare la qualità delle transazioni, il reporting, le analisi di business, permettendo nel contempo di ridurre i costi per la protezione delle informazioni e i rischi derivanti dalla loro perdita, furto o distruzione.

2.2.7 Data Integration and Interoperability

Considerato che la maggior parte delle organizzazioni dispone di centinaia o migliaia di database e archivi, la gestione dei processi di trasferimento dei dati tra i vari storage di dati all'interno dell'organizzazione (ma anche per altre organizzazioni) è diventata di fondamentale importanza. Il processo di trasferimento e combinazione dei dati, se non gestito correttamente, può sovraccaricare le risorse e le capacità IT. Dalla necessità di gestire la complessità ed i costi ad essa associati nasce quindi il bisogno nella prospettiva aziendale di integrare i dati. Una buona progettazione aziendale per l'integrazione nativa di dati è molto più efficiente ed economica rispetto ad altre soluzioni come point-to-point o distribuite, il cui sviluppo può comportare un elevato numero di interfacce di comunicazione tra le differenti applicazioni e mettere in seria difficoltà anche le organizzazioni con IT più capaci ed organizzate. Gli "hub" di dati come data warehouse e Master Data permettono di ridurre questo problema, consolidando i dati necessari alle applicazioni e fornendo viste coerenti dei dati stessi. Allo stesso modo, la complessità di gestione di dati operativi, che sono condivisi in tutta l'organizzazione, può essere notevolmente ridotta attraverso tecniche di integrazione dei dati aziendali come l'integrazione "hub-and-spoke".

Un aspetto fondamentale per tenere sotto controllo i costi delle operations aziendali è la gestione dei costi di supporto. Questi possono infatti aumentare quando i dati vengono trasferiti utilizzando più tecnologie, ciascuna delle quali richiede competenze specifiche per lo sviluppo e la manutenzione. Accorgimenti come l'identificazione di tool standard possono ridurre i costi di supporto e del personale e migliorare l'efficienza delle attività di analisi dei problemi su un prodotto o processo aziendale. Ridurre la complessità di gestione delle interfacce può ridurre i costi di manutenzione delle stesse e permettere una distribuzione più efficiente delle risorse di supporto aziendali, indirizzandole su altre attività organizzative.

Un caso significativo di gestione della Data Integration è il processo base ETL di Estrazione, Conversione e Caricamento del dato (Extract, Transform and Load) usato in particolare per la gestione dei Data Warehouse. Il processo di estrazione include la selezione dei dati necessari e la loro estrazione dalla fonte di origine. Dopo essere stati estratti i dati vengono organizzati in uno store fisico definito target. Eventuali tecniche di trasformazione dei dati li rendono compatibili con la struttura del data store target. Questi dati dovrebbero essere pronti per l'integrazione con i dati nella struttura target. Il caricamento dei dati nel processo ETL rappresenta la memorizzazione o presentazione fisica del risultato delle trasformazioni nel sistema target. A questo punto i dati, in

base alle trasformazioni eseguite, allo scopo del sistema target e all'uso previsto, potrebbero richiedere un'ulteriore elaborazione o essere pronti per essere presentati agli utilizzatori.

Lo sviluppo tecnologico ha permesso di ottenere soluzioni di integrazione dei dati in grado di garantire un allineamento estremamente veloce dei dati tra due database distinti. Queste tecnologie sono richieste in settori di business che necessitano di un controllo e di un monitoraggio "real-time" o comunque a bassa latenza di determinati fenomeni e possono comportare costi elevati in hardware e software specializzati. Le soluzioni di integrazione di dati a bassa latenza sono progettate per ridurre al minimo i tempi di risposta agli eventi come nel caso di aggiornamenti continui riguardo all'acquisto di beni o titoli finanziari, commenti sui social media e letture da sensori che monitorano posizione, temperatura, utilizzo o altri valori.

Tra le tecniche di integrazione esistenti, una delle più comunemente utilizzate è l'utilizzo di un modello canonico, che standardizza i formati in cui i dati sono condivisi. Ad esempio in pattern di progettazione d'integrazione dati hub-to-spoke, tutti i sistemi che vogliono ricevere o fornire dati interagiscono con un solo hub informativo centrale. I dati vengono trasformati da o verso un sistema di invio o ricezione basato su un messaggio in formato standard, ovvero su un modello canonico. Sebbene sviluppare e concordare un formato standard condiviso sia un obiettivo impegnativo da raggiungere, avere un modello canonico può ridurre in modo significativo la complessità dell'integrazione e interoperabilità dei dati in azienda e per tanto, ridurre notevolmente i costi di supporto. La creazione e la gestione di un modello di dati canonico comune per tutte le iterazioni di dati è un costo necessario nell'implementazione di una soluzione di integrazione di dati aziendali.

Un aspetto da presidiare nei processi di data integration articolati, che coinvolgono sistemi lungo una filiera di trattamento dei dati è il data lineage ovvero l'insieme dei singoli passaggi del percorso che i dati attraversano. Il data lineage è utile per determinare l'origine di un dato a valle di una filiera, derivarne la conseguente affidabilità in base alle fonti, analizzare l'impatto dovuto a modifiche di sistema o problemi legati ai dati. In ogni caso, affinché possa essere documentata la conoscenza delle origini e dello spostamento dei dati, è necessaria la Data Governance. In un mondo dove i dati vengono sempre più condivisi anche tra aziende, esistono limitazioni d'uso e standard di conformità emergenti, come ad esempio, per il settore assicurativo, il regolamento Solvency II in Europa, che impone alle organizzazioni di mantenere chiaro il percorso dei propri dati, da dove provengono e come sono stati modificati nel trasferimento attraverso vari sistemi. Per misurare la

portata e i vantaggi dell'implementazione delle soluzioni di integrazione di dati, è opportuno considerare metriche su disponibilità, volume, velocità, costi e utilizzo. Alcune metriche di riferimento possono essere i costi di sviluppo e gestione delle soluzioni, la facilità di acquisizione di nuovi dati, la complessità di soluzioni e operazioni, oppure il numero di sistemi che utilizzano soluzioni di integrazione dei dati. In ogni caso per una buona soluzione di integrazione dati sono fondamentali le regole di business che permettono di definire cosa ci si aspetta dai processi di business stessi, anche perchè è importante sottolineare come siano regole e requisiti di business, e non scelte tecnologiche, determinare i requisiti di un processo di integrazione dei dati.

Principali problemi di integrazione

Data l'importanza della fase di integrazione e la sua complessità, è necessario analizzare i possibili problemi che potrebbero nascere da questo processo. Uno dei problemi maggiori per l'integrazione di dati è sicuramente la grande quantità di dati provenienti da fonti differenti. Questo risulta un problema sia per l'archiviazione dei dati stessi, sia per l'estrapolazione di informazioni utili dall'analisi di questa enorme mole di dati. Non sempre per le imprese risulta immediata la scelta di quali dati raccogliere, considerato che essi potrebbero essere utili in quel momento oppure potrebbero esserlo in un futuro più o meno prossimo, anche perché lo storage di dati rappresenta un costo, che cresce al crescere della quantità di dati raccolti. Il secondo problema ricorrente nel mondo dei dati nasce dal fatto che si debbano integrare dati da numerose fonti differenti. Sfida ampiamente riconosciuta in vari settori, da quello bancario o assicurativo a quello energetico e manifatturiero. Per garantire un uso ed integrazione ottimale dei dati, è opportuno bisognerebbe prima di tutto fare una valutazione su quali dati siano veramente necessari per l'organizzazione, cercando così di prioritizzarli, evitando così di accumulare grosse quantità di dati senza un fine di utilizzo chiaro. Inoltre, è assolutamente necessario comprendere quali siano gli utilizzatori finali dei dati e capirne i bisogni in modo da mantenere efficiente il processo. Il terzo problema è costituito dall'integrazione manuale di dati. Molte imprese si affidano a Data Scientist, i quali spendono molto del loro tempo per raccogliere ed elaborare i dati. Questa operazione potrebbe essere svolta in maniera più veloce ed efficiente tramite un processo automatizzato di integrazione dati, svolto da una delle molte piattaforme presenti sul mercato. Questo permette di avere moltissimi benefici, tra cui ad esempio: l'integrazione e relativa documentazione dei dati, il posizionamento dei dati in maniera tale da uniformarli sotto gli stessi standard, la lettura di chiavi comuni che potrebbero collegare i dati per successive analisi, raggruppandoli per campi. Infine, la possibilità di tracciare il

data lineage, avendo a disposizione lo storico di tutti gli step evolutivi del dato, in maniera tale da potersi muovere sul processo di evoluzione del dato per ulteriori analisi. Un altro problema di notevole importanza è rappresentato dalla presenza di diverse tecnologie dei sistemi sorgente di dati. Occorre infatti valutare la tipologia di database che viene utilizzata in azienda e verificare la possibilità di svolgere attività operative sui dati presenti nel database stesso.

2.2.8 Data Warehousing and Business Intelligence

Il concetto di Data Warehouse nacque attorno agli anni 80', quando la tecnologia iniziò a permettere alle organizzazioni di integrare i dati provenienti da varie fonti in un ambiente integrato di consultazione, esplorazione, analisi. Gli obiettivi principali di questo ambiente sono il supporto, delle attività di analisi, delle informazioni e dei processi decisionali. Il Data Warehouse generalmente è costituito da un database relazionale per archiviare e gestire i dati e da soluzioni software come ETL, che permettono di estrarre, trasformare e caricare i dati per prepararli alle successive analisi. Un esempio di problema comune per le organizzazioni è quello di dover fornire sempre più spesso dati che dimostrino di aver soddisfatto i requisiti normativi e, dal momento che contengono uno storico più o meno aggiornato, i Warehouse sono spesso il mezzo per rispondere a queste richieste. Dal momento che i DW riuniscono dati che saranno utilizzati in modi diversi e per vari scopi, e poiché il loro utilizzo si evolve nel tempo man mano che gli utenti analizzano ed esplorano i dati, è fondamentale un'accurata progettazione, che se ben organizzata, permette un notevole risparmio in costi di rielaborazione. Nella raccolta dei requisiti è quindi opportuno concentrarsi sugli obiettivi e la strategia di business, cercando di individuare le metriche e indicatori chiave per la misurazione e il monitoraggio dei fenomeni di interesse dell'azienda.

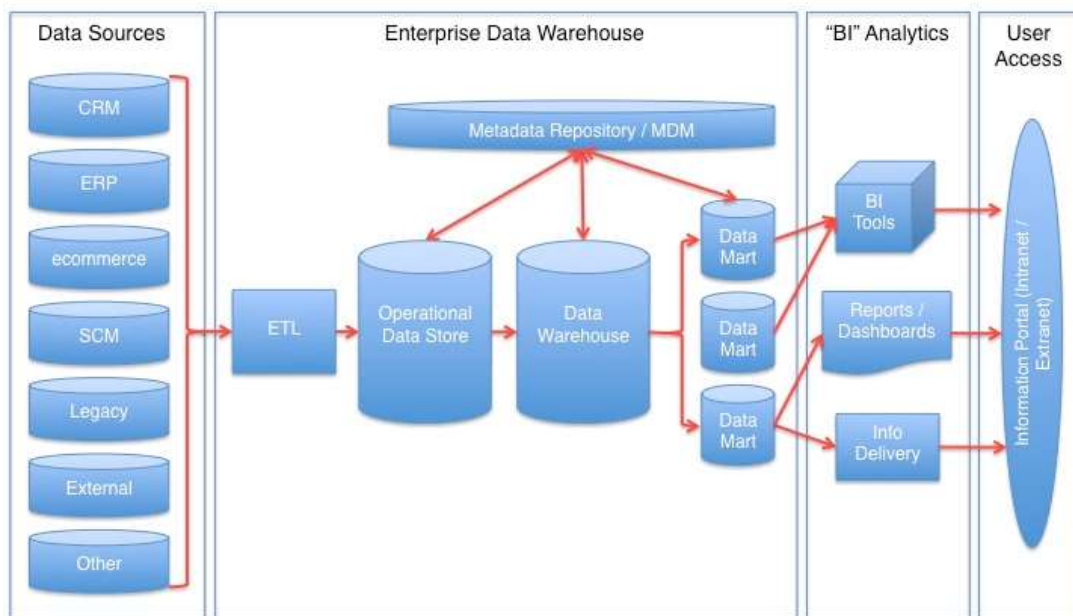


Figura 4 Struttura Data Warehouse

Il DW descritto da Bill Inmon, Claudia Imhoff e Ryan Souasa nel contesto di Corporate Information Factory (CIF) include varie fasi e componenti. Tra di essi ad esempio:

- I dati che provengono da più fonti, come macchinari differenti in un sito produttivo, vengono trasformati attraverso strumenti di ETL, in modo da poter essere integrati con la rappresentazione/modello aziendale standard nel Data Warehouse.
- I dati fluiscono successivamente all'interno di un'Operational Data Storage (ODS): si tratta di un database integrato di dati operativi. Questi tipi di database generalmente contengono dati correnti o a breve termine (30-90 giorni), mentre un DW contiene anche dati storici (spesso vengono raccolti storici di dati aziendali di diversi anni). Un ODS può servire come fonte primaria per un DW oppure può essere usato per controllarlo. Inoltre, dal momento che un ODS contiene una finestra temporale di dati e non lo storico, può essere aggiornato e modificato molto più rapidamente di un warehouse.
- I dati vengono memorizzati all'interno del database relazionale del DW.

- Per svolgere analisi più rapide i dati possono essere presentati sotto forma di Data Mart (Datamarts), che spesso rappresentano un sottoinsieme di DW progettato per supportare un particolare tipo di analisi o un gruppo specifico di consumatori di dati.
- A questo punto i dati sono pronti per le attività di Business Intelligence, in modo tale da procedere con analisi statistiche, reporting e data mining. Esistono strumenti che permettono di presentare e visualizzare i dati agli utenti aziendali. Per le attività di Business Intelligence è fondamentale conoscere gli utilizzatori finali dei dati in modo da permettere di fornire loro il livello di interattività con il dato che necessitano. Ad esempio un supply chain manager potrebbe aver bisogno di analizzare un report statico sui dati finanziari, ma anche un report molto interattivo per l'analisi dell'inventario. Quindi è necessario permettere a manager e altri utenti di poter identificare le cause primarie dei problemi fino anche, se necessario, all'analisi del singolo dato. Tra le migliori pratiche per il monitoraggio e la messa a punto della BI vi è la definizione e la visualizzazione di un insieme di metriche per la valutazione del soddisfacimento dei clienti. Per analisi in "real-time" o di breve periodo si può accedere ai dati anche direttamente dagli ODS, senza passare necessariamente dal Data Warehouse.

Il Data Warehouse rappresenta il punto di integrazione dei dati aziendali per supportare il processo decisionale del management, l'analisi strategica e la pianificazione di attività e risorse. I dati che necessitano di una correzione possono essere scartati oppure corretti alla fonte e poi nuovamente alimentati all'interno del sistema. Considerando inoltre aspetti legati al ciclo di vita dei dati, questi si spostano dai sistemi sorgente in un'area di staging dove possono essere puliti e arricchiti man mano che vengono integrati e memorizzati nel DW oppure all'interno di un ODS. Non è sempre necessaria la costruzione di un Data Warehouse e, solo dopo una attenta analisi si può verificare la sua reale utilità all'interno di un contesto aziendale. Ad esempio se per assurdo l'esigenza aziendale fosse quella di supportare una singola analisi, che porti alla determinazione di un unico insieme di indicatori, non sarebbe necessario impostare un processo di concentrazione dei dati in un ambiente di DW per poi renderli disponibili per le finalità dell'analisi stessa. La costruzione di un Data Warehouse comincia a convenire quando sono richieste numerose analisi. In questo caso infatti, si potrebbero trarre benefici dal DW sia in termini di efficienza che in termini di consistenza dei dati stessi. Infatti è probabile che per diverse analisi servano gli stessi dati, che avranno quindi subito lavorazioni tali da avere gli stessi standard.

Il supporto di Business intelligence rimane l'altro aspetto principale per avere all'interno dell'azienda un DW. La BI fornisce una visione dell'organizzazione, dei suoi clienti e dei suoi prodotti. Inoltre, un'organizzazione che agisce sulla base delle conoscenze acquisite dalla BI può migliorare l'efficienza operativa ed il vantaggio competitivo. Con il passare del tempo, il miglioramento della qualità e delle quantità di dati, una maggiore velocità di elaborazione degli stessi, la BI si è evoluta da una valutazione retrospettiva ad una analisi predittiva.

2.2.9 Data Quality

La fase di valutazione della qualità dei dati dovrebbe sempre essere svolta in un contesto di Data Management. Migliore è la qualità dei dati, migliori saranno i risultati delle analisi. Naturalmente le organizzazioni possiedono una grandissima quantità di dati, ma non tutti hanno la stessa importanza. Cercare di mantenere la qualità dei dati elevata è uno sforzo ben giustificato dal fatto che il costo per isolare i dati giusti la prima volta è minore rispetto a isolare quelli sbagliati e poi applicare azioni correttive successivamente. Inoltre, stabilire criteri di Data Quality all'inizio di un processo è un chiaro segno di organizzazione matura di Data Management. Naturalmente non tutti i processi necessitano della stessa qualità dei dati, questa può essere definita come "fit for purpose", ovvero commisurata al reale bisogno. Non è sempre necessario infatti definire a priori un livello di qualità dei dati molto elevato quando non richiesto, basta che siano soddisfatti i requisiti. Alla base di un *Data Quality Management Cycle* vi è l'identificazione dei dati che non rispondono ai requisiti di qualità degli utilizzatori e i conseguenti problemi e rischi connessi a queste non conformità. Questa analisi permette di identificare le cause del problema e capire quanto costi risolverli. Il ciclo nel dettaglio è rappresentato dalla struttura Plan-Do-Check-Act.

Per concentrarsi sulla qualità dei dati ci sono una serie di attività chiave da svolgere:

1. Definire il grado di qualità dei dati: per fare questo ci sono una serie di domande che permettono di capire come migliorare e settare il giusto livello di qualità dei dati.
2. Definire una strategia di Data Quality: per analizzare le azioni e le politiche di Data Quality è inoltre necessario utilizzare metriche di valutazione, in modo tale da poter informare gli utilizzatori dei dati sulle caratteristiche di qualità che sono importanti per il loro uso.

3. Identificare dati critici e regole di business: i dati non hanno tutti lo stesso valore. Se si punta su una maggiore qualità nei dati che sono più importanti, l'azienda riuscirà a dare più valore all'organizzazione e ai clienti.
4. Performare una valutazione iniziale di Data Quality: da qui si possono avere notevoli miglioramenti. Si usa quindi un piccolo effort concentrato per dimostrare come funzioni il processo di miglioramento.
5. Identificare e definire priorità a eventuali miglioramenti dopo aver dimostrato che il processo di miglioramento ed implementazione può funzionare.
6. Definire obiettivi per i miglioramenti di Data Quality: deve esserci un ritorno positivo nell'investimento in miglioramento della qualità dei dati. Ad esempio, una volta riconosciuti i problemi, si può calcolare il ROI in seguito al miglioramento della qualità dei dati.
7. Sviluppo e spiegazione delle DQ operations: ad esempio mettere in atto un programma che permetta al team di gestire le DQ rules, gli standard e altri elementi.

2.2.10 Metadata

La definizione più comunemente utilizzata per descrivere i metadata è "data about data". Per comprendere meglio questo concetto si consideri una fotografia registrata su un supporto elettronico: questa, oltre alla mera immagine contiene una serie di informazioni principali come ora e data in cui è stata scattata, formato, dispositivo dal quale è stata scattata. Queste caratteristiche sono i metadati della fotografia. In ogni caso le entità a cui sono applicabili i metadati sono numerose: flussi informativi, dataset, campi relativi, e per estensione processi che li trattano e relative unità organizzative impegnate nella loro gestione, regole e vincoli di business e tecnici. Ognuna di queste entità può essere corredata da un insieme specifico di metadati che ne rappresentano le caratteristiche. I metadati aiutano quindi un'organizzazione a comprendere i suoi dati, sistemi e flussi di lavoro, permette inoltre di valutare la qualità dei dati ed è parte integrante per la gestione di database e applicazioni. I metadati possono essere quindi considerati come un "catalogo" che permette di classificare e qualificare i dati e di reperirli, insieme al loro corredo descrittivo attraverso diversi criteri di ricerca. Per fare un esempio, un'organizzazione senza metadati può essere considerata come una biblioteca senza un catalogo. All'interno di un'organizzazione viaggiano su vari livelli e direzioni moltissimi dati, e nessuno potrà conoscere

tutto su di essi, rischiando quindi di perdere preziosa conoscenza e vantaggio competitivo. Per questo motivo risulta fondamentale avere uno strumento come i metadati che permetta di documentare le informazioni, catturare e gestire le conoscenze organizzative sui dati.

Esistono tre tipologie principali di metadati: di business, tecnici e operazionali. I metadati di business si concentrano principalmente sul contenuto e condizioni dei dati, includendo dettagli relativi alla governace dei dati. I metadati tecnici forniscono informazioni sui dettagli tecnici dei dati, sui sistemi che li archiviano e sui processi che li muovono tra i vari sistemi. I metadati operazionali descrivono in maniera più dettagliata il trattamento e l'accesso ai dati.

2.2.11 Reference e Master Data

Generalmente con Reference Data si intendono i dati che vengono utilizzati per caratterizzare, classificare altri dati o per correlare dati ad informazioni esterne all'organizzazione, quindi Reference Data Management rappresenta il controllo e la manutenzione di determinati valori di dominio, definizioni e relazioni all'interno e tra i valori di dominio. L'obiettivo è garantire che i valori siano coerenti e attuali tra le varie funzioni facendo in modo che i dati siano accessibili all'organizzazione. Anche i Reference Data richiedono metadati per permettere che la loro definizione sia efficace. Sono tipicamente informazioni stabili con un determinato set di valori che cambia raramente e non dipende dalle scelte interne all'organizzazione, ma alle quali dovrà adattarsi di fronte a variazioni e modifiche dei Reference Data. Un esempio è rappresentato dalla lista di province all'interno di uno stato.

Con Master Data si intendono invece entità di business definite all'interno dell'organizzazione come dipendenti, clienti, prodotti, strutture finanziarie, asset, che forniscono il contesto per le transazioni e le analisi di business. I Master Data rappresentano i dati più accurati disponibili sulle entità di business chiave e, se gestiti correttamente, i valori dei Master Data possono essere considerati attendibili ed utilizzati in sicurezza. Il Master Data Mangement rappresenta il controllo su valori di Master Data e su identificativi che permettono di avere un uso coerente dei dati più precisi e puntuali sulle entità di business.

La definizione di Gartner del Master Data Management sottolinea come sia una disciplina composta da persone, processi e tecnologia e non semplicemente una soluzione applicativa specifica. Egli sostiene che il Master Data Management sia infatti "una disciplina compatibile con la tecnologia in cui business ed IT lavorano insieme per assicurare uniformità, precisione, stewardship, coerenza

semantica e affidabilità degli asset dei Master Data ufficiali condivisi dell'azienda. I Master Data sono l'insieme coerente e uniforme di identificatori e attributi estesi che descrivono le entità centrali dell'impresa, tra cui clienti, clienti potenziali, cittadini, fornitori, siti, gerarchie e piano dei conti".

2.3 Come esercitare il Data Management in pratica

Il concetto sul quale bisogna concentrarsi nel momento in cui poi effettivamente si deve esercitare il Data Management all'interno di un'organizzazione ruota attorno all'espressione "planning for better data". Con questa frase si intende la creazione di un processo strutturato che permetta di gestire i modi attraverso cui i dati si collegano ai processi di business e allo stesso tempo identificare i dati più importanti, che permettano di raggiungere gli obiettivi aziendali strategici di lungo periodo. Occorre sempre tenere a mente che spesso i dati si muovono in maniera orizzontale all'interno di un'organizzazione, ma gli obiettivi di business risultano invece verticali, rendendo quindi indispensabile un punto di incontro tra queste due diverse dimensioni, anche considerando il fatto che i dati siano gestiti in modi diversi, da soggetti diversi e nascano in luoghi diversi. Per questi motivi è ancora più importante costruire una tassonomia a livello aziendale attraverso Governance e Master Data in modo che ci siano delle definizioni univoche all'interno dell'organizzazione. Un processo di Data Management comporta importanti cambiamenti all'interno di un'impresa: organizzativi ma anche culturali che impongono quindi di non cercare subito di "boil the ocean", ovvero gestire ogni singolo dato aziendale, creando probabilmente in questo modo confusione e risultati scadenti. Una strategia vincente invece si concentra all'inizio su pochi dati, cercando innanzitutto di identificare i dati critici e iniziare il processo di Data Management proprio con questi. Inoltre è necessario, per un discorso di qualità dei dati, scartare i dati obsoleti, incompleti o appunto di bassa qualità dal momento che rappresentano un rischio, portando con sé un'informazione distorta e molto probabilmente errata. Da qui si definisce quindi la strategia attraverso la Data Governance con un piano contenente la descrizione di tutte le attività, quando verranno svolte e da chi. Successivamente, assieme all'IT, verranno definiti i modi per realizzare questo piano di azione strategico.

3 DATI E MERCATI

Come considerato fino ad ora, il Data Management è una disciplina che coinvolge diverse aree aziendali, diverse figure ed uno sforzo più o meno accentuato in base agli obiettivi di business che muovono la gestione dei dati. Occorre ora concentrarsi sui possibili settori di applicazione del Data Management. Ci sono moltissimi esempi di come attraverso l'utilizzo degli strumenti tecnologici attuali si possano estrarre più informazioni dai dati raccolti da un'organizzazione. Per i prodotti di consumo si utilizzano analisi per ottenere conoscenza dalla spesa di marketing e dai dati sulle prestazioni di vendita, con conseguente programmi di marketing più mirati. Nelle attività bancarie vengono applicate tecniche quantitative per comprendere le correlazioni tra prestazioni aziendali, condizioni macroeconomiche e risultati interni. Per l'organizzazione interna e gestione delle risorse le analisi sono volte a migliorare la redditività in tutta l'azienda, consentendo ai leader di individuare ciò che realmente guida le prestazioni. Infine sempre nel settore finanziario l'elaborazione dei pagamenti sfrutta l'apprendimento automatico per individuare nuovi modelli e formulare regole per automatizzare la categorizzazione dell'allocazione del denaro. Si considerino ora con maggiore attenzione alcuni di questi settori, in particolare finanziario, energetico e manifatturiero, molto diversi tra loro, ma tutti accomunati dall'esigenza di una gestione efficiente delle grandi quantità di dati che producono e scambiano.

3.1 Data Management nel settore finanziario

In generale il settore finanziario comprende per definizione tutte le attività connesse ad operazioni finanziarie come ad esempio gestione degli investimenti, accesso al credito, transazioni sui mercati finanziari, servizi di consulenza finanziaria ma anche controllo della gestione finanziaria in una società o delle finanze pubbliche. Le imprese principalmente coinvolte in questo settore sono i cosiddetti intermediari finanziari come ad esempio banche, che possono essere sia commerciali che di investimento, oppure compagnie di assicurazioni, ma anche altre società come operatori di borsa e fondi di investimento, banche centrali e società di consulenza finanziaria.

Questo settore è caratterizzato da un elevatissimo numero di transazioni e operazioni che si verificano con altissima frequenza, che coinvolgono diversi attori e clienti, generando quindi una quantità enorme di informazioni sotto forma di dati. All'interno del settore finanziario inoltre, i controlli legislativi e burocratici sono molto severi, riducendo il margine di errore e imponendo alle

aziende che si trovano in questo settore di mantenere un controllo maggiore sui dati. Questi sono solo alcuni dei motivi per cui il Data Management risulta fondamentale nel settore finanziario, dove il miglioramento dei processi di acquisizione, gestione dei dati e rendicontazione finanziaria, possono avere un effetto immediato sui principali dati operativi e sui modelli di rating. Una maggiore automazione di questi processi aziendali porta ad un aumento dell'efficienza e di conseguenza a un più alto livello di redditività per le aziende coinvolte. Ma l'automazione richiede un livello più elevato di qualità dei dati, quindi la necessità di una governance professionale dei dati. Completato dall'introduzione di un processo di Data Security che consenta una gestione efficace dei dati anagrafici del cliente e una sua conseguente conoscenza precisa e sicura.

I Big Data sono un paradigma emergente in quasi tutti i settori e i big data finanziari (FBD) stanno diventando una delle aree di gestione e governance più promettenti nel settore finanziario. Molti ricercatori sostengono che i Big Data stiano alimentando la trasformazione della finanza e in generale del business in modi che possono ancora essere valutabili. Una nuova area di ricerca si sta evolvendo per studiare modelli quantitativi e approcci econometrici per gli studi finanziari che possono colmare il divario tra la ricerca finanziaria empirica e la scienza dei dati. In questa affascinante area, esperti e scienziati possono proporre nuovi modelli di business finanziario utilizzando i Big Data, presentare metodi sofisticati per il controllo del rischio con strumenti di machine learning, fornire strumenti di visualizzazione per l'analisi dei mercati finanziari. Senza contare poi la possibilità di creare nuovi indici di sentiment finanziario estraendo i sentimenti del pubblico da i massicci dati testuali dai social network e distribuire gli strumenti basati sulle informazioni in altri modi creativi.

Analizzando più da vicino qualche esempio pratico, anche nei fondi di investimento i dati stanno assumendo un ruolo sempre più chiave. L'ultimo decennio ha visto infatti ampie riforme patrimoniali e regolamentari, con misure finali come quelle derivanti dal framework normativo Basilea e un approccio al rischio di credito rivisto per il portafoglio bancario, che dovrebbe essere implementato nei prossimi anni. Tutto ciò mira a rafforzare la resilienza alle flessioni e agli shock di mercato, semplificando i bilanci delle banche, aumentando i livelli di capitale e applicando un trattamento punitivo al capitale e alla liquidità per le attività finanziarie che mostrano bassi livelli di trasparenza. Poiché sia le società di vendita che quelle di acquisto continuano ad aumentare le loro esposizioni ai fondi, è importante comprendere i dati sottostanti e le sfide tecnologiche affrontate dagli investitori quando soddisfano requisiti di rischio e normativi sempre più rigorosi.

Con la crescente importanza dei fondi come veicolo di investimento per ottenere rendimenti e diversificazione, l'accesso a dati di alta qualità non significa solo costi aggiuntivi per soddisfare i requisiti di reporting normativo. Ci sono chiari vantaggi aziendali e di gestione del rischio da ottenere. Le banche possono ridurre le riserve di capitale di rischio e le compagnie di assicurazione hanno una visione chiara della performance dei fondi nel loro universo di investimento, consentendo l'allocazione basata sui dati e le decisioni di selezione del gestore. Per avere successo oggi, le istituzioni finanziarie devono rispondere a sfide non affrontate dagli approcci tradizionali. Le aziende devono affrontare la sfida dei dati sui fondi in modo strategico, puntando su soluzioni di livello mondiale che integrino processi, dati e sistemi di rischio.

Un esempio di proposta nel mondo finanziario è quella di Rongcai Hu, Meng Liu, Pingping He e Yong Ma, in "Can Investors on P2P Lending Platforms Identify Default Risk?", che decidono di utilizzare i dati di per esaminare la relazione tra i tassi di interesse sui prestiti e i rischi di insolvenza dei mutuatari. Attraverso questa analisi, scoprono che esiste un'asimmetria tra il tasso di interesse e il rischio di insolvenza. Cioè, gli ordini con lo stesso tasso di interesse possono avere diversi rischi di insolvenza. I loro risultati empirici mostrano anche come, nonostante quello che ci si possa aspettare, maggiore è il reddito di un mutuatario, maggiore è il rischio di insolvenza. Inoltre, gli autori ritengono che gli investitori possano commettere errori sulla relazione tra alcune informazioni di un mutuatario e il rischio di insolvenza, ma potrebbero identificare il rischio di insolvenza del mutuatario purché abbiano una buona conoscenza dei principali fattori che influenzano il rischio di insolvenza, come ad esempio l'importo, la durata e il livello di rischio di credito del mutuatario.

Un altro esempio è rappresentato dal documento "Come si influenzano l'un l'altro i mercati azionari globali? Evidence from Finance Big Data and Granger Causality Directed Network ", dove Yong Tang, Jason Jie Xiong, Yong Luo e Yi-Cheng Zhang applicano il test di causalità Granger per costruire il Granger Causality Directed Network per 33 principali indici del mercato azionario globale per studiare come i mercati si influenzano a vicenda indagando i margini diretti nelle diverse reti filtrate. I risultati rivelano che le classifiche sono simili in entrambi gli approcci in cui gli indici degli Stati Uniti dominano la prima posizione seguiti da altri indici americani, europei e asiatici. L'analisi del lead-lag rivela che ci sono effetti di lag tra gli indici globali. Il risultato getta nuove intuizioni sulle influenze tra i mercati azionari globali con implicazioni per la progettazione della strategia di trading, la gestione del portafoglio globale, la gestione del rischio e la regolamentazione dei mercati.

3.2 Data Management nel settore energetico

Con il rapido sviluppo della tecnologia dei sensori, di trasmissione wireless, di comunicazione in rete, del cloud computing e dei dispositivi mobili intelligenti, grandi quantità di dati vengono accumulate durante quasi tutti gli aspetti di vita quotidiana. Il volume dei dati sta crescendo rapidamente con strutture e forme sempre più complesse. In particolar modo nel settore energetico vengono generate grandi quantità di dati sulla produzione e il consumo di energia e si digitalizzano i sistemi energetici, con la crescente penetrazione delle tecnologie dell'informazione emergenti. Le aziende energetiche stanno cercando a tal proposito, di applicare tecnologie innovative per ridurre i margini di errore, competere con una forte concorrenza e lavorare intorno a rigide normative governative. Attraverso una gestione efficiente dei dati su diversi fronti come adeguato supporto IT, governance, integrazione, elaborazione, analisi e sicurezza, le società energetiche possono aumentare i loro profitti e ridurre gli sprechi, e i consumatori disporre di energia quando ne hanno bisogno. I dati infatti permettono di scoprire nuove fonti di energia, ridurre i costi di perforazione ed esplorazione, aumentare l'efficienza dell'esplorazione e la produttività, prevenire e prevedere incidenti durante l'utilizzo di macchinari come trivelle per l'estrazione del petrolio. Inoltre garantiscono l'assenza di interruzioni di corrente, promettendo agli utenti una distribuzione dell'energia 24 ore su 24. Un'analisi volta a valutare come le persone consumano energia permette di regolare l'offerta in base all'andamento delle previsioni di domanda. Infine è possibile consentire un risparmio energetico nelle case e nelle industrie e prevedere processi di manutenzione e riparazione.

Vi sono numerosi esempi di gestione efficace dei dati nel settore energetico. Uno tra questi è rappresentato dalla sismologia. Le aziende energetiche utilizzano infatti il monitoraggio sismico, o indagini sismiche, per scoprire nuovi terreni ricchi di petrolio. Utilizzando strumenti di monitoraggio le società energetiche sono in grado di valutare la redditività del petrolio appena scoperto, grazie anche a numerose piattaforme di elaborazione parallele. L'analisi dei dati viene ampiamente utilizzata per stimare la quantità di petrolio o gas che deve ancora essere estratta da un pozzo, determinare la qualità del suolo, indagare su eventuali anomalie geologiche e confrontare i dati storici di produzione, la storia delle trivellazioni petrolifere locali, il clima e modifiche dell'ambiente, in modo tale da facilitare una migliore perforazione e produzione. Anche l'estrazione del petrolio stessa è diventata un processo high-tech. Si consideri un giacimento petrolifero in cui ogni macchina è in comunicazione con le altre e tutte trasmettono costantemente dati alla sede centrale dell'azienda energetica. Le piattaforme petrolifere sono in comunicazione con i loro operatori per

mantenere un flusso di produzione costante. I sensori avvisano i minatori di eventuali riparazioni necessarie e i compressori avvisano i lavoratori ogni volta che si verifica un sovraccarico pericoloso. In linea con queste idee, l'azienda petrolifera Chevron utilizza una tecnologia intelligente per migliorare i processi di estrazione e produzione di petrolio. BP ha lo stesso tipo di tecnologia e la chiama Field-of-the-Future, mentre Royal Dutch Shell utilizza la tecnologia Smart Field, costituita da giacimenti petroliferi digitalizzati che raccolgono tutti i dati. Le tecnologie utilizzate da queste aziende collezionano inoltre i dati dei sensori tra cui temperatura, pressione, volume, vibrazioni e urti, combinano i dati con analisi e comunicazioni internazionali in tempo reale per consentire alle società energetiche di monitorare ogni fase del processo di produzione del petrolio. Questi passaggi includono il momento in cui il petrolio viene scoperto, l'estrazione, la manutenzione quotidiana del pozzo e della macchina, la purificazione e la distribuzione. Tramite raccolta, integrazione e gestione dei dati efficiente, le aziende energetiche possono trarre profitto da ogni goccia di petrolio estratta riducendo i costi. Le compagnie che operano nel settore del gas e del petrolio hanno notato inoltre come, incorporando la gestione dei dati nella gestione dell'energia, sia possibile rilevare e prevenire potenziali incidenti prima che si verifichino. Utilizzando sensori, le aziende possono monitorare il ciclo di vita di diversi componenti e macchinari, e sostituirli prima che causino un disastro. Quando ad esempio la temperatura e la pressione superano un valore soglia critico, viene trasmesso un avviso al centro di controllo e vengono subito attuate misure di sicurezza. Se il terreno su cui poggia il pozzo è fragile e non è in grado di sostenere la perforazione, questo viene rilevato e il processo si interrompe. Infine, quando le macchine richiedono la sostituzione di parti, attraverso la trasmissione di informazioni, gli operatori possono avere le parti pronte prima che queste siano completamente inutilizzabili e causino un blocco del processo.

Oltre all'estrazione del petrolio e al mantenimento della sicurezza nell'industria petrolifera, il Data Management viene utilizzato anche nel settore delle energie rinnovabili. Ad esempio, durante la raccolta dell'energia eolica, le turbine ruotano e possono essere utilizzate per raccogliere dati sul vento. Queste informazioni vengono utilizzate per monitorare le prestazioni delle turbine, inclusi eventuali guasti. Inoltre, i dati sul vento combinati con i dati di funzionamento della turbina, le vibrazioni e le emissioni acustiche, i dati degli ordini di lavoro come i dati delle riparazioni e la storia della turbina come gli anni di produttività, mostrano all'azienda numerosi ambiti in cui migliorare l'efficienza e ridurre i costi. Proprio per la produzione di energia rinnovabile i dati meteorologici, come l'angolo del sole, la velocità del vento e la temperatura, svolgono un ruolo importante nel supportare la gestione intelligente dell'energia. Ad esempio, i dati meteorologici possono essere

utilizzati per la previsione della produzione di energia da fonti rinnovabili e la previsione del consumo energetico degli utenti, supportando così i processi decisionali dei diversi partecipanti ai sistemi energetici. Questo risulta molto interessante considerando il fatto che la maggior parte dei consumatori, compresi i privati cittadini e le industrie, spreca energia per il modo in cui la usa. Pertanto, attraverso un'attenta analisi del consumo di energia si possono ridurre i consumi invece di aumentare la produzione. Proprio per questo obiettivo comune, le aziende energetiche e altri settori stanno lavorando per ridurre lo spreco di energia. Sistemi domestici come Google Home, termometri Nest e sensori stanno già riducendo il consumo di energia domestica. Il National Institute of Science and Technology prevede che entro il 2030 le aziende utilizzeranno l'analisi dei dati delle reti intelligenti per risparmiare fino a 2 trilioni di dollari. Poiché il petrolio sta diventando più scarso e man mano che l'energia pulita diventa più diffusa, vi è una chiara necessità di incorporare efficacemente i dati da tutte le aree della produzione di energia. Dall'estrazione al consumo, il Data Management nella gestione dell'energia è un elemento chiave per rendere l'energia disponibile, redditizia e sicura.

Un segno importante del fatto che esista una crescita di interesse nel settore della gestione intelligente dell'energia basata sui dati, è rappresentato dalla costante nascita di nuove start-up legate a queste tematiche. I prodotti forniti da queste start-up sono principalmente legati a cloud computing, analisi dei big data e tecnologie di rilevamento basate su dispositivi hardware, software, piattaforme e sistemi intelligenti. I loro servizi e soluzioni si concentrano principalmente su monitoraggio e previsione in tempo reale, risposta alla domanda e gestione dal lato della domanda, coinvolgimento dei clienti, ottimizzazione dell'efficienza energetica, notifiche e rapporti sui consumi energetici e marketing mirato.

Dal momento che i dati energetici vengono raccolti da molte fonti diverse, utilizzando differenti dispositivi di acquisizione dati, l'integrazione diventa un aspetto molto importante per la loro analisi. I dati provenienti da diverse piattaforme e applicazioni sono generalmente eterogenei, indipendenti e reciprocamente chiusi. Inoltre, la struttura, il formato e la qualità dei dati variano notevolmente. Oltretutto, i rischi e la privacy sono questioni chiave durante l'intero processo di gestione intelligente dell'energia basata sui dati. Per ottenere vantaggi economici e sociali derivanti dai dati prodotti nel settore energetico, la privacy degli individui deve essere efficacemente protetta e i potenziali rischi dell'utilizzo dei dati vanno ragionevolmente evitati.

3.3 Data Management nel settore manifatturiero

Il settore manifatturiero continua ad essere un elemento fondamentale del motore economico mondiale, ma la sua posizione di fronte al cambiamento tecnologico è cambiata drasticamente. Per questo motivo anche in questo settore sono necessari nuovi metodi per guidare la produttività, aumentare l'efficienza e l'innovazione. Proprio a tal fine i dati e una loro gestione efficiente può rappresentare un elemento fondamentale per evitare che il cambiamento possa essere disruptive, permettendo alle imprese di scoprire nuove informazioni e identificare modelli che consentano loro di migliorare i processi, aumentare l'efficienza della catena di approvvigionamento e isolare le variabili che influenzano la produzione. I leader delle aziende manifatturiere comprendono la posta in gioco. Uno studio di Honeywell Process Solutions-KRC Research ha evidenziato come il 67% dei dirigenti di produzione prevede di investire nella raccolta e analisi dei dati, anche di fronte alle pressioni per ridurre i costi. La maggioranza capisce che queste analisi sono necessarie per competere con successo in un'economia basata sui dati e sta facendo per questo motivo investimenti nell'integrazione dei dati e nelle risorse di gestione per realizzare una corretta trasformazione digitale e ottenere vantaggio competitivo.

L'Internet of Things (IoT) ha aggiunto una nuova dimensione con asset e sensori connessi in grado di misurare, registrare e trasmettere le informazioni sotto forma di dati in tempo reale. Questi dati sono potenzialmente di grande valore per i produttori, ma molti sono sopraffatti dall'enorme volume di informazioni in arrivo. L'utilizzo di corrette procedure di Data Management può aiutare le aziende manifatturiere a raccogliere, ripulire, integrare e analizzare i dati provenienti dai macchinari per rivelare informazioni che possono contribuire al miglioramento delle prestazioni. Alcuni esempi risultano particolarmente significativi, come quello di un importante produttore chimico europeo concentrato sul miglioramento della resa. Utilizzando i dati provenienti dai sensori, è stato analizzato il modo in cui ogni fattore di input influenzava l'output di produzione. L'analisi è proseguita quindi con temperature, quantità, flusso di anidride carbonica e pressioni del refrigerante e sono stati confrontati i tassi di influenza tra loro. Di conseguenza, hanno rilevato che le portate di anidride carbonica influenzano enormemente la resa e modificando leggermente i parametri. Il risultato è stato una significativa riduzione degli sprechi di materie prime (del 20%) e dei costi energetici (del 15%), migliorando notevolmente la resa. Un altro esempio di miglioramento della resa è rappresentato da una grande azienda farmaceutica. Anche in questo caso sono stati analizzati i dati dei sensori delle apparecchiature, rivelando le interdipendenze tra i vari parametri

di produzione di vaccini e infine è stato considerato il modo in cui ciascuno di essi ha influenzato la resa. Quindi, i 9 parametri più cruciali sono stati identificati, esaminati e regolati per ottimizzare il processo di produzione, questo ha permesso di migliorare la resa dei vaccini del 50%. Infine si consideri un'azienda produttrice di zucchero, con un problema di qualità all'interno della propria produzione. L'azienda rilevava elevati livelli di umidità e materie prime di bassa qualità, che hanno gravemente influenzato il gusto dello prodotto. Per combatterlo, sono stati raccolti e analizzati i dati dei sensori per cercare delle correlazioni tra i parametri che contribuiscono alla migliore qualità dello zucchero. Con questa intuizione il produttore è riuscito a trovare un modo per influenzare rapidamente la qualità del prodotto e raggiungere uno standard di zucchero unificato indipendentemente da fattori esterni. Questo approccio ha permesso di ridurre i costi di produzione, aumentare la soddisfazione dei clienti e semplificare i carichi di lavoro.

I dati, oltre a consentire l'analisi di uno storico, possono guidare l'analisi predittiva, che i produttori possono utilizzare ad esempio per pianificare la manutenzione, consentendo così alle aziende di prevenire costose interruzioni degli asset ed evitare tempi di fermo imprevisti. Un esempio è rappresentato dalle apparecchiature di fabbrica dell'azienda Intel, le quali trasmettono in streaming i dati generati dall'IoT. Questi dati vengono utilizzati per il riconoscimento dei modelli e il rilevamento dei guasti, consentendo di vedere quali tendenze richiedono attenzione immediata e quali azioni sono necessarie per prevenire gravi guasti in officina. Tale manutenzione predittiva permette di ridurre in maniera drastica il tempo di reazione e di ridurre i costi. Anche il noto produttore automobilistico Rolls-Royce utilizza ampiamente dati nella sua produzione. Una delle loro esperienze di gestione di dati più interessanti è connessa alla modellazione di nuovi motori aeronautici. In fase di progettazione, il loro software, integrato con strumenti di Data Management, crea simulazioni di nuovi motori a reazione e analizza terabyte di dati per vedere se i nuovi modelli siano validi. Ciò consente all'azienda di trovare i punti deboli prima che il modello arrivi alla produzione, il che riduce i costi legati ai difetti e aiuta a progettare il prodotto di una qualità molto superiore.

La gestione e analisi dei dati può avere un impatto significativo: lo studio Honeywell-KRC ha scoperto che attraverso queste analisi è possibile ridurre i guasti fino al 26% e tagliare i tempi di inattività non programmati di quasi un quarto. Con le giuste analisi, i produttori possono concentrarsi su ogni segmento del processo di produzione ed esaminare le catene di approvvigionamento nei minimi dettagli, tenendo conto delle singole attività. Questa capacità di restringere il focus consente ai

produttori di identificare i colli di bottiglia e rivelare processi e componenti con prestazioni insufficienti. Per esempio un'azienda che necessita un modo migliore per gestire la consegna delle materie prime, in modo tale da evitare i costi legati ai fallimenti della catena di approvvigionamento. Utilizzando i dati, compatibilmente con i vincoli di privacy, contenenti i dettagli del percorso dei loro fornitori, nonché i dati meteorologici e sul traffico forniti da fonti esterne affidabili nasce la possibilità di identificare la probabilità di ritardi nelle consegne. Per fare ciò, attraverso gestione ed analisi di questa enorme quantità di dati provenienti da fonti differenti, è possibile generare analisi predittive, calcolando possibili ritardi e carenze di materie prime. Infine, sulla base di questi calcoli, l'impresa può elaborare un piano di emergenza relativo alla fornitura e sarà in grado di eseguire la produzione senza interruzioni ed evitare costi eccessivi per i tempi di fermo.

Un processo strutturato di Data Management permette alle imprese manifatturiere di rivelare le dipendenze, consentendo ai produttori di migliorare i processi di produzione e creare piani alternativi per affrontare potenziali insidie, prevedendo inoltre con precisione la domanda di prodotti personalizzati. Attraverso l'analisi dei dati che contengono informazioni relative al cambiamenti nel comportamento dei clienti, i produttori possono avere più tempo di consegna, offrendo l'opportunità di produrre personalizzazioni con la stessa efficienza quasi quanto i beni prodotti su scala più ampia. Alcune funzionalità innovative includono strumenti che consentono agli ingegneri di prodotto di raccogliere, analizzare e visualizzare il feedback dei clienti quasi in tempo reale. A tal proposito un esempio lampante è rappresentato da un cliente di ScienceSoft del settore automobilistico, che utilizza i dati per fornire supporto post-vendita ai propri utenti e garantire un miglioramento continuo. I centri operativi del cliente analizzano in tempo reale grandi quantità di dati provenienti dai sensori dell'auto (dati diagnostici, chilometraggio, geolocalizzazione, ecc.) e generano approfondimenti sulle prestazioni del prodotto. L'analisi di questi dati consente all'azienda di monitorare lo stato del prodotto, annotare e persino prevedere alcuni malfunzionamenti, offrendo un servizio di manutenzione preventiva volto ad evitare danni più gravi. Tale approccio consente al cliente di aumentare la qualità del prodotto e migliorare l'esperienza del consumatore finale.

Infine, fornendo ai produttori gli strumenti di cui hanno bisogno per approfondire i processi, l'analisi dei dati consente loro di identificare i punti all'interno del processo di produzione in cui inserire in modo proficuo processi personalizzati utilizzando capacità interne o posticipando la produzione per consentire a un partner di eseguire la personalizzazione in anticipo al completamento del processo

di fabbricazione. Secondo un'analisi Deloitte sull'aumento della personalizzazione di massa, la possibilità di posticipare la produzione offre ai produttori una nuova flessibilità che consente loro di accettare richieste su ordinazione. Deloitte osserva inoltre che la capacità di posticipare la produzione può "aiutare a ridurre i livelli di inventario e, in ultima analisi, aumentare l'efficienza dell'impianto". Un processo di produzione semplificato non è solo vantaggioso di per sé, ma offre ai produttori un modo per mantenere l'efficienza durante l'esecuzione delle personalizzazioni.

4 IL CASO REALE DELL'AZIENDA BRUGNAGO S.R.L.

4.1 L'azienda

L'azienda Brugnago nasce nel 1933 come selleria per la produzione di borse porta attrezzi per autoveicoli civili, industriali e macchine di movimento terra. L'azienda nel corso degli anni si evolve e sposta il proprio business sulla progettazione e produzione meccanica di utensili di soccorso, con la conseguente commercializzazione dei kit completi alle case automobilistiche. Questa nuova linea di business consente all'azienda di crescere, fino a diventare, negli anni 80', leader indiscusso nel mercato italiano. Brugnago continua ad evolversi, investendo nel 1996 in una nuova linea di produzione con macchinari CNC tecnologicamente avanzati, ampliando i servizi di progettazione e incrementando il controllo qualità sulla produzione. Nel 2007 nasce il gruppo Brugnago tramite l'acquisizione dell'azienda Huashen Automobile Tools con sede a Shanghai, leader cinese della fabbricazione di dotazioni utensili per automotive. Integrità, eccellenza, esplorazione ed efficienza sono soltanto alcuni dei valori cardine che hanno consentito a Brugnago questa evoluzione, e che permettono di definirla ai giorni nostri come Industria 4.0.

La sede italiana di Brugnago, sulla quale si svilupperà il caso studio della tesi, conta 44 dipendenti, suddivisi in 30 operai specializzati e 14 impiegati, e un fatturato medio annuale pari a 10-12 milioni di euro. L'azienda negli ultimi anni ha puntato molto sulla digitalizzazione. In particolare è stato sviluppato un progetto di implementazione del sistema MES (Manufacturing Execution System), attraverso la realizzazione di un applicativo software, denominato Overone, per la programmazione e l'avanzamento della produzione. In questo modo ora è possibile monitorare in real-time lo stato di avanzamento degli ordini e inviare feedback istantanei su eventuali malfunzionamenti dei macchinari. L'obiettivo principale dei progetti all'interno di Brugnago è la digitalizzazione del dato di produzione, delle attività di pianificazione e delle registrazioni dei fermi produttivi. Attraverso l'introduzione di strumenti software come Overone all'interno dell'azienda, il grado di digitalizzazione e il numero di dati prodotti e scambiati è cresciuto. Da qui nasce il focus della tesi, che si pone come obiettivo la comprensione di come tutti questi dati generati dagli impianti, possano essere utilizzati da Brugnago per ottenere vantaggio competitivo e quindi valore. Verranno quindi analizzate le tecniche di gestione e integrazione dei dati, e le problematiche che possono sorgere per raggiungere l'obiettivo prefissato.

4.2 I dati dell'azienda

Per procedere con una valutazione delle soluzioni per la gestione dei dati da introdurre all'interno di Brugnago, è necessario comprendere quali siano le caratteristiche principali dei dati che sono generati e condivisi dall'azienda. In particolare le analisi sui dati hanno portato alle seguenti evidenze:

- Attraverso lo scambio di dati tra il software Overone installato a bordo dei macchinari di produzione e i server aziendali, è stato possibile stimare il numero di dati scambiati in un giorno all'interno dell'azienda dal reparto produttivo. Il calcolo è stato svolto su un macchinario, monitorando il flusso di dati in upload e download per un'intera giornata di produzione. Grazie a questo calcolo è emerso che i dati prodotti in 24 ore all'interno del reparto produttivo sono circa 1.12 GB. Considerando che all'interno dello stabilimento sono presenti 20 macchinari collegati al software Overone, è possibile stimare che in un'intera giornata di produzione possono essere scambiati fino a 22.4 GB di dati.
- I dati all'interno dell'azienda sono di due tipologie principali. I dati di produzione e programmazione della produzione e i dati di controllo qualità. I dati di produzione sono ad esempio codice dell'ordine, quantità richiesta in un determinato ordine, descrizione delle fasi di produzione, codice macchina di lavorazione, tempi di lavorazione e di fermo. Le basi dati sul controllo di qualità contengono invece informazioni quali codice dell'ordine, caratteristica analizzata, tolleranza, scostamento dalla tolleranza, valore di scostamento.
- Analizzando meglio la struttura dei record di dati in analisi, è stato possibile evidenziare come all'interno siano presenti molti campi ripetuti. Questo dipende dal modo in cui l'azienda, in collaborazione con i fornitori dei vari software, abbia deciso di sviluppare le varie query e dal modo in cui vengono estratti i dati. Considerando ad esempio un file di export del software Overone, questo contiene varie colonne, che rappresentano informazioni quali:
 1. Codice ordine di produzione
 2. Nome operatore
 3. Codice del macchinario
 4. Ora inizio e ora di fine
 5. Causale di fermo

Il file è impostato in modo che molti campi, come numero di macchinario e codice dell'ordine, si ripeteranno lungo tutto il record. In questo caso a venire memorizzati non saranno i campi ripetuti

più volte all'interno del record, ma il singolo dato di riferimento nel dataset. Il dataset contiene varie tabelle che includono ad esempio:

- I codici degli ordini di produzione
- I codici dei macchinari
- I nomi degli operatori
- I nomi dei controlli di qualità che possono essere svolti

Per questo motivo il numero di dati che vengono memorizzati è minore del numero di dati scambiati giornalmente dal server con i software aziendali. Il numero di dati memorizzati a sistema dipende infatti dal numero di entità presenti, come il numero di codici di ordine e il numero di macchinari.

- Gli standard dei dati all'interno dell'azienda sono di due tipi diversi:
 1. Standard di formato: all'interno dei dataset sono presenti più formati come stringhe, numeri interi o floating point, date, valori numerici. A questi corrispondono per esempio nome dell'operatore, data di inizio e fine produzione, valori di scostamento durante il controllo qualità.
 2. Standard legati al protocollo dei singoli macchinari: in questo caso le informazioni sono registrate in formati distinti nei diversi macchinari, che seguono protocolli differenti. È possibile infatti che le macchine possiedano standard proprietari in funzione dei produttori del macchinario.

In Brugnago i dati non hanno il secondo problema di formato, in quanto il software Overone, durante la raccolta dati dai diversi macchinari di produzione, li trascrive in modo che abbiano lo stesso standard. Nel caso in cui l'azienda non possedesse uno strumento come Overone che garantisce già una standardizzazione dei dati da diversi macchinari, sarebbe necessario introdurre un processo di integrazione. Quest'ultimo deve contenere al suo interno uno strumento che permetta di svolgere una trascodifica, in modo che i dati provenienti dai macchinari di produzione con standard differenti, siano convertiti nello standard del sistema di destinazione. Il problema può nascere, oltre che dalla differenza di protocolli, anche dall'ordinamento dei dati. Le colonne che costituiscono la struttura del record possono contenere le stesse informazioni, ma in ordine differente, anche in questo caso è necessario

specificare al sistema di destinazione non solo il formato, ma anche l'ordine delle informazioni che verranno ricevute.

Per svolgere un'integrazione e standardizzazione, quando sono presenti in azienda macchinari che producono dati con formati proprietari differenti, è necessario un numero di operazioni più o meno elevato e complesso in funzione di diversi parametri, come:

- numero dei macchinari
- numero degli step di integrazione
- protocolli proprietari dei sistemi sorgente di dati

4.3 Il problema

L'azienda Brugnago ed in generale molte aziende nel settore manifatturiero, che presentano un grado medio-alto di digitalizzazione, non comprendono come gestire ed estrarre valore dai dati prodotti, dal momento che esistono delle difficoltà per lavorare con i dati in azienda. I problemi principali, derivanti dal basso e mancato utilizzo dei dati in azienda, sono dovuti a più motivazioni:

- All'interno di un'azienda digitalizzata ma non "data driven", non sono presenti figure e competenze legate alla gestione dei dati, che comporta la mancanza di un team per la gestione ed analisi dei dati.
- Dal momento che mancano le competenze legate alla gestione dei dati, mancano anche procedure consolidate di governance dei dati e anche una cultura aziendale focalizzata sul dato e sulla sua gestione.
- Non sono presenti e quindi non utilizzati strumenti e software di data analytics.

4.4 Proposta di soluzione

La tesi propone due soluzioni di gestione ed integrazione dei dati, le quali non sono esclusive tra di loro, ma sequenziali.

- La prima soluzione è rappresentata dall'automatizzazione dei processi di raccolta ed integrazione dati per il monitoraggio dell'indice Overall Equipment Effectiveness, che attualmente viene calcolato manualmente dagli operatori. È stata scelta questa proposta di soluzione perchè mostra come un intervento con investimento limitato in termini di tempo, costi e risorse, possa migliorare notevolmente la qualità dei dati aziendali. L'azienda

Brugnago infatti, attraverso una revisione dei processi che governano il flusso dei dati e un'integrazione automatizzata dal software Overone, può creare una base dati integrata di valore. Quest'ultima rappresenta un tassello fondamentale per la gestione dei dati all'interno dell'azienda e per lo sviluppo di successive analisi.

- La seconda soluzione consiste nello sviluppo di un modello, attraverso la piattaforma di data management Irion EDM, sviluppata dall'azienda software torinese Irion. Questa soluzione nasce dal fatto che le analisi svolte sul caso studio dell'azienda Brugnago dimostrano come, in generale, all'interno di aziende del settore manifatturiero, non sono presenti competenze e strumenti per il supporto alla gestione dei dati. Introdurre uno strumento come quello proposto da Irion può essere una soluzione a queste carenze. L'utilizzo di uno strumento software specifico all'interno dell'azienda permette infatti di ottenere una governance migliore dei dati e di un monitoraggio migliore da parte del management aziendale. In generale, attraverso la soluzione si evidenzia un modo per utilizzare i dati prodotti dalla digitalizzazione degli impianti nelle imprese manifatturiere.

4.5 Soluzione 1: automatizzazione dei processi di raccolta ed integrazione dati per il monitoraggio dell'indice Overall Equipment Effectiveness

4.5.1 L'indice Overall Equipment Effectiveness OEE

Un tema comune che spesso si trovano ad affrontare le aziende, in particolare quelle nel settore manifatturiero, è rappresentato dal miglioramento dell'efficienza dello stabilimento aziendale. Questo miglioramento viene inteso non come miglioramento puntuale in base alla risposta di singole esigenze, ma in un miglioramento continuo delle prestazioni dell'impianto. In quest'ottica di miglioramento, si possono trovare alcune chiavi comuni di lettura per il miglioramento dell'efficienza di impianto:

- Gli obiettivi strategici e operativi devono confrontarsi sempre con l'impatto sugli impianti (sales and operation planning)
- La qualità di un determinato prodotto o servizio è strettamente legata alle logiche di gestione dell'impianto
- Le competenze di un impianto, seguendo quindi una logica bottom-up, sono un driver fondamentale per il miglioramento dell'efficienza di tutta l'azienda

In questo contesto spicca in particolar modo un indicatore concentrato proprio sull'impianto: l'Overall Equipment Effectiveness (OEE). Quest'ultimo è un indicatore standard nel panorama internazionale, che si concentra sulla misurazione della produttività aziendale. In particolare, misura il tempo durante cui un impianto è effettivamente produttivo. L'OEE è una grandezza adimensionale che può assumere valori tra 0 e 1, spesso espressa in punti percentuali e viene calcolata per un determinato intervallo di tempo t , come ad esempio un turno di 8 ore in azienda. Esso riassume in sé tre diversi aspetti molto importanti dal punto di vista del controllo di produzione: la disponibilità, la resa e la qualità del processo. In generale il calcolo è il seguente:

$$\text{OEE} = \text{Disponibilità} \times \text{Resa} \times \text{Qualità}$$

La disponibilità viene calcolata come il rapporto tra il tempo operativo e il tempo disponibile per la produzione. Il tempo disponibile per la produzione è il tempo totale disponibile della macchina (quindi tutto l'anno) meno le chiusure pianificate, come le manutenzioni programmate, le manutenzioni preventive, mancanza di ordini, chiusure aziendali. Il tempo operativo è il tempo disponibile per la produzione meno il tempo in cui la macchina è occupata per attività non pianificate superiori a qualche minuto, come rotture, setup, cambi versione, aggiustamenti. La disponibilità tiene conto dei tempi di stop programmato e non programmato dell'impianto, se avessimo disponibilità del 100% vorrebbe dire che il processo si svolge soltanto durante il tempo di produzione programmato.

La resa valuta le perdite di performance, considerando tutti i fattori che possono far muovere il processo manifatturiero più lento della massima velocità possibile. Alcune possibili cause di perdita di velocità e quindi di performance sono: usura dei macchinari, materiali di lavorazione scadenti ed inceppamenti. La performance viene calcolata come il rapporto tra il tempo di produzione effettivo a consuntivo a fine turno al netto delle fermate, e il tempo previsto di produzione nel turno. La performance serve per misurare quanto l'impianto venga utilizzato in funzione della velocità dei cicli di produzione, cicli più lenti e brevi fermi macchina possono rallentare la performance di impianto. Performance del 100% significa che il processo viene svolto il più velocemente possibile.

Il fattore qualità considera la perdita di qualità, considerando le parti prodotte che non rispondono agli standard di qualità. Alcuni esempi di bassa qualità sono gli elementi di scarto dalla produzione oppure parti che presentano difetti di vario tipo e per questo devono essere rilavorati. Per valutare la qualità dell'impianto, vengono conteggiati i pezzi buoni prodotti rispetto ai pezzi totali prodotti,

potendo così ottenere le perdite dell'impianto. Da questa informazione è possibile dedurre la perdita di tempo durante la produzione, legata alla scarsa qualità dei pezzi prodotti non conformi alle specifiche. Un impianto che offre un livello di qualità del 100%, non produce pezzi difettosi durante di cicli produttivi.

Il calcolo dell'OEE non migliora automaticamente la produttività, ma per un'azienda manifatturiera è importante calcolarlo, in quanto consente di avere una misura dell'efficienza all'interno dell'impianto. A partire dal valore ottenuto è possibile valutare i parametri di miglioramento, andando a lavorare sulle cause che provocano perdita di efficienza e quindi di valore. Il monitoraggio dell'indice OEE, consente inoltre di valutare come questo migliori o peggiori a valle di progetti ed azioni di miglioramento messe in atto. La natura sistemica dell'OEE permette all'azienda di monitorare tre diverse variabili, sulle quali è possibile focalizzarsi di volta in volta, concentrando maggiormente gli sforzi su quello che, tra i tre indici considerati, risulta il più basso. Un valore obiettivo di tutto rispetto della grandezza OEE si aggira intorno all'intervallo 85-88%. Il Japan Institute of Plant Maintenance (JIPM), l'inventore di questo parametro, ogni anno consegna il TPM awards in cui vengono premiate tra le altre, le aziende che hanno raggiunto un valore dell'OEE piuttosto elevato, pari a 88-90%.

Questo indicatore risulta importante per l'analisi della tesi, dal momento che è costituito a sua volta da tre diversi indicatori, che raccolgono dati da fonti diverse all'interno dell'azienda. Nasce quindi la possibilità di un'integrazione di questi dati automatizzata e la necessità di procedure di data governance. In questo modo è possibile avere una gestione ed un tracciamento affidabili di grandi quantità di dati prodotti in azienda che contribuiscono alla creazione di risultati precisi, ottenuti attraverso informazioni affidabili, che garantiscano analisi di valore per l'azienda.

4.5.2 Fonti di dati

Considerando un processo di integrazione dati è importante comprendere quante siano le fonti di dati all'interno dell'azienda per ottenere le informazioni necessarie. Nel caso specifico le informazioni necessarie sono relative al calcolo dell'indice OEE. Queste sono tutte informazioni a cui un'azienda manifatturiera dovrebbe avere accesso, dal momento che i dati di maggiore importanza sono relativi alla programmazione di produzione, alla produzione e all'analisi della qualità dei pezzi prodotti. Il grande lavoro di digitalizzazione svolto all'interno dell'azienda Brugnago, ha consentito di ridurre il numero di fonti di dati in azienda, migliorandone

contemporaneamente l'affidabilità. Le fonti dati principali all'interno dell'azienda considerate per l'analisi sono tre: il sistema gestionale ERP aziendale denominato DNA, il software di controllo di produzione Overone e il macchinario CMM per il controllo qualità.

L'ERP DNA rappresenta il cuore dell'azienda. Contiene tutte le informazioni relative alla programmazione della produzione, agli ordini e al loro stato di avanzamento. In particolare, considerando gli ordini di produzione attualmente abilitati, è possibile visualizzare contemporaneamente tutti gli ordini e le loro informazioni che sono stati rilasciati e pronti per essere lavorati all'interno dell'azienda.

Sui macchinari di produzione è montato un tablet connesso ad un software denominato Overone. Questo rappresenta la seconda importante fonte di dati nell'impianto. L'ERP aziendale DNA e il software Overone comunicano tra di loro. Attraverso una semplice procedura: i dati relativi agli ordini di produzione pronti per essere processati vengono esportati dall'ERP aziendale al software di gestione delle macchine Overone. Una volta fatta questa prima operazione è possibile vedere lo storico degli ordini. Su Overone si possono trovare informazioni importanti come il codice dell'ordine in fase di lavorazione, macchinari coinvolti, descrizione articolo, codice cliente, ecc... Sia il gestionale DNA che il software Overone conoscono le fasi di lavorazione di un determinato ordine, incluse eventuali lavorazioni che non devono essere svolte all'interno dello stabilimento. Ad esempio se un pezzo dovesse essere verificato esternamente, oppure subire lavorazioni esterne all'azienda, questa informazione sarebbe presente. Il software Overone contiene inoltre informazioni importanti per il calcolo dell'OEE relative alle manutenzioni programmate, che derivano dall'ERP aziendale, e informazioni relative ai fermi non programmati in produzione. Queste ultime sono particolarmente importanti per il calcolo dell'OEE, dal momento che spesso, gran parte dell'inefficienza in un processo produttivo nasce proprio da microfermate e fermi non programmati. A tal proposito, nella sezione "analisi tempi" di Overone vengono mostrati esattamente i tempi di lavorazione e i tempi di fermo, con le relative causali. Questi possono essere ad esempio un'ora per assemblea, mezz'ora per la pausa pranzo, o altre microfermate in modo tale che i tablet a bordo macchina di Overone rilevino in automatico il tempo di lavorazione e il tempo di fermo per ogni ordine di lavorazione. A fine di ogni turno ci sarà inoltre il numero di pezzi lavorati, confrontabile con i pezzi teorici, che si possono trovare sul gestionale DNA in modo tale da fare una valutazione della performance.

La terza fonte di dati è rappresentata dalla macchina di misura a coordinate CMM, utilizzata per il monitoraggio della qualità della produzione in azienda. Durante la lavorazione vengono effettuati controlli di qualità sui prodotti, queste informazioni vengono poi aggiornate su Overone come valutazione dei pezzi scartati durante l'ordine di produzione. I dati prodotti durante il controllo qualità al momento non sono integrati né con il gestionale DNA, né con il software Overone. Vengono infatti raccolti manualmente all'interno di un record excel dagli operatori addetti ai controlli qualità.

Gli investimenti sostenuti da parte dell'azienda per la digitalizzazione hanno permesso di adottare soluzioni come il software Overone molto importanti per il monitoraggio della produzione nello stabilimento. Senza uno strumento di questo tipo infatti, la raccolta dei dati dovrebbe essere svolta manualmente dagli operatori di reparto. Questo porterebbe, oltre ad un grande investimento in termini di tempo da parte degli operatori, ad un risultato limitante, con una elevata probabilità di commettere errori, scarsa affidabilità delle fonti e quindi conseguente bassa qualità dei dati raccolti. L'utilizzo di dati di questo tipo all'interno dell'azienda per il calcolo di indici come l'OEE oppure per altre analisi, comporta rischi ed inefficienze, e potrebbe addirittura peggiorare le analisi svolte in ottica di miglioramento, compromettendo le decisioni operative e strategiche che verranno prese successivamente. Attraverso uno strumento di raccolta dati automatizzata come il software Overone, le informazioni sono chiare e facilmente accessibili. La qualità dei dati raccolti risulta notevolmente maggiore rispetto ad una raccolta svolta manualmente. Considerando un turno di lavoro, il software Overone registra ogni volta che il macchinario si ferma, con tempistiche precise al secondo, per un operatore sarebbe alquanto difficile ottenere lo stesso risultato. Uno strumento di questo tipo porta con sé quindi notevoli benefici, ma anche costi di investimento, giustificati però dalla effettiva differenza nella qualità dei risultati delle analisi svolte e dalle possibilità di gestione dei dati in maniera più semplice e comprensibile. Questa soluzione rappresenta l'unico modo per monitorare gli impianti, avendo un flusso di dati e quindi di informazioni di qualità, provenienti da una fonte verificata ed affidabile.

Un aspetto importante per mantenere la qualità dei dati elevata per il calcolo dell'indice OEE risulta anche la collaborazione umana da parte degli operatori. Il software infatti rileva con elevata precisione i tempi di lavorazione e di fermo, ma è l'operatore che deve causalizzare i fermi e i ritardi di produzione. Queste causali risultano importanti per le successive analisi sulle cause di fermo e di rallentamento sulla linea. Per questo motivo è necessaria un'attenta formazione sul funzionamento

del software in modo tale da poter operare nel miglior modo possibile e generare dati il più accurati e quindi di alta qualità possibile.

Il caso studio affrontato nella tesi che coinvolge l'azienda Brugnago rappresenta quindi una situazione già favorevole, con un numero limitato di fonti di dati. In generale nelle aziende manifatturiere la situazione potrebbe essere diversa e i dati provenire da molte più fonti. In tal caso sarebbero necessari degli interventi in azienda volti a introdurre uno strumento di raccolta e standardizzazione dei dati da diverse fonti, ad esempio un software di raccolta dati come Overone, in modo da rendere più semplice la gestione dei dati e le successive analisi.

4.5.3 Situazione attuale per il monitoraggio dell'indice OEE

Attualmente l'indice OEE viene già monitorato dall'azienda Brugnago mediante un rapporto semplificato di produzione elaborato attraverso un file excel. All'interno del file vengono riportati manualmente dagli operatori i dati principali per il calcolo dell'OEE. Viene calcolata prima la resa attraverso le informazioni che derivano dall'ERP sui valori programmati di lavorazione, sia in termini di pezzi previsti che di tempo previsto di lavoro e dal software Overone per valutare tempo e pezzi effettivi. In questo modo viene monitorato il tempo che un operatore avrebbe dovuto impiegare per produrre un determinato numero di pezzi e il tempo reale che impiega per la produzione. La valutazione prosegue poi con le perdite di tempo imputabili rispettivamente a disponibilità, performance e qualità e si conclude con il calcolo dell'OEE risultante dato dal prodotto delle tre componenti. La soluzione attuale di calcolo e monitoraggio della produzione permettere di avere informazioni sufficienti per gestire la produzione. Uno strumento con un maggiore affidabilità e privo degli errori compiuti con questo metodo darebbe spazio ad analisi più precise ed una maggiore implementazione.

4.5.4 Problemi legati alla situazione attuale

Il problema principale di questo metodo di integrazione manuale dei dati per il calcolo dell'OEE, è il fatto che sia svolto manualmente dagli operatori. Questo infatti porta delle conseguenze non trascurabili per le successive analisi:

- Bassa qualità dei dati utilizzati per le analisi
- Fonti di dati con affidabilità limitata, dal momento che vengono raccolti ed integrati manualmente dagli operatori all'interno del file excel

- Calcolo dell'OEE e analisi successive (ad esempio delle causali di fermo) con bassa accuratezza, che possono comportare errori e quindi decisioni imprecise da parte del management.
- Il report generato dagli operatori in base ai calcoli effettuati viene generato in maniera soggettiva e non oggettiva

4.5.5 Implementazione dell'integrazione automatizzata

La soluzione di integrazione dati proposta per il caso dell'azienda Brugnago prevede un'integrazione automatizzata dei dati dalle principali fonti. Queste come detto in precedenza sono tre: il sistema gestionale ERP aziendale DNA, il software di monitoraggio della produzione Overone installato a bordo dei macchinari e il macchinario di controllo qualità CMM. Lo sviluppo della soluzione è articolato in tre step principali:

1. Nel caso non fosse ancora presente in azienda, introduzione all'interno della manifattura di uno strumento di monitoraggio della produzione digitalizzato. All'interno di Brugnago, tramite un progetto di digitalizzazione del MES, è già stato introdotto un software di raccolta dati di produzione connesso ai macchinari, denominato Overone. Quest'ultimo permette così di evitare che i dati di produzione siano raccolti manualmente dagli operatori, ma bensì da uno strumento automatizzato.
2. Nel caso dell'azienda Brugnago la successiva implementazione è rappresentata dall'integrazione del macchinario di controllo qualità CMM, con il software Overone. Questa operazione consentirebbe di avere a disposizione sul software Overone sia le informazioni di produzione che quelle di qualità dei prodotti, permettendo così di risparmiare sui tempi, ridurre i rischi e gli errori. Questa integrazione è piuttosto semplice in termini di tempi e fattibilità, l'aspetto che però non va tralasciato in termini organizzativi all'interno dell'azienda è rappresentato dalla formazione del personale addetto al controllo qualità. È necessario infatti procedere con un'attenta formazione sulla registrazione dei dati non in forma manuale su un file excel ma in automatico sul software Overone, il quale prevede un'area dedicata al controllo qualità, nella quale inserire soltanto i dati utili dal macchinario. Superata la formazione, il beneficio sarebbe notevole, anche in termini di causali di fermo macchina. Se ad esempio la macchina CMM si surriscaldasse, questo potrebbe essere classificato per causa dall'operatore sul sistema Overone, dopo aver registrato in automatico il tempo esatto di fermo.
3. Il successivo sviluppo è rappresentato dall'integrazione dei dati in maniera automatica direttamente dal software di monitoraggio della produzione, rispetto all'integrazione manuale

dei dati sul file excel da parte degli operatori, che avviene al momento per il calcolo dell'indice OEE.

4.5.6 Situazione dopo l'implementazione

Grazie a queste implementazioni i dati verrebbero raccolti dai macchinari attraverso il software Overone e le classificazioni per cause verrebbero fatte dagli operatori di turno, annotando semplicemente sul software il motivo del fermo accanto al tempo corrispondente (ad esempio 30 minuti per pausa pranzo, oppure 1 ora per rottura e cambio inserto). A questo punto i dati necessari per il calcolo dell'OEE sarebbero trasferiti automaticamente dal software Overone e dal Gestionale aziendale in un foglio di calcolo, per poi procedere con il calcolo dell'indice o di altri indici di produttività e resa utili per il reparto. A valle dei valori ottenuti il manager potrà fare valutazioni sulla produzione, sul lavoro degli operatori e prendere decisioni sugli eventuali interventi, considerando anche le giuste priorità.

Tabella 1 Cambiamento del processo di raccolta e integrazione dati

	Prima dell'automatizzazione	Dopo l'automatizzazione
Raccolta dati	Operazione svolta manualmente dagli operatori	Operazione svolta dal software Overone
Integrazione/ trasferimento	Operazione svolta manualmente dagli operatori	Operazione svolta dal software Overone

4.5.7 Benefici

I benefici delle due soluzioni sono stati categorizzati in base alla possibilità di poterli quantificare oppure soltanto qualificare, e se sia possibile farlo prima dell'introduzione della soluzione o dopo. In caso sia possibile quantificarli è stata descritta una proposta di metodo per stimare il beneficio.

Qualificabili ma non quantificabili

- Minore trasferimento manuale e manipolazione dei dati da parte degli operatori
- Sviluppo contenuto di nuove competenze, limitato semplicemente alla formazione sui software utilizzati in reparto

Quantificabili prima dell'adozione della soluzione

- Risparmio di tempo da parte degli operatori nella compilazione dei file excel, con conseguente possibilità di utilizzo degli operatori per altre operazioni.

Quantificabile attraverso il calcolo del tempo che mediamente impiega un operatore per la raccolta e integrazione manuale dei dati all'interno dei fogli di calcolo.

Quantificabili dopo l'adozione della soluzione

- Qualità dei dati utilizzati più elevata.

Quantificabile attraverso un confronto con la situazione attuale relativamente ad alcuni parametri tra cui ad esempio: numero di dati scorretti, numero di dati ridondanti, numero di dati ripetuti.

- Maggiore affidabilità delle fonti di dati: i dati vengono trasferiti in una base dati integrata direttamente dal software di raccolta dati di produzione e dal gestionale aziendale.

Quantificabile attraverso un confronto con la situazione attuale relativamente ad alcuni parametri tra cui ad esempio: numero di dati scorretti, numero di dati ridondanti, numero di dati ripetuti.

- Definizione e contestualizzazione migliore delle cause che possono far nascere problemi, ritardi e fermi macchina. L'indice OEE risulta particolarmente utile per l'azienda, in quanto permette di identificare le problematiche che incidono maggiormente sull'inefficienza produttiva. In base al risultato ottenuto, le perdite precedentemente classificate per causa vengono ordinate dalla maggiore alla minore in termini di dimensione e frequenza, per poter lavorare sulle cause dei problemi che sono emersi. È interessante infatti valutare, attraverso opportuni grafici, la distribuzione delle cause di fermo per ogni macchina, osservare in questo modo i fenomeni che si ripetono con maggiore frequenza e generare automaticamente una priorità nelle problematiche su cui agire. Nell'esempio mostrato in figura è rappresentata l'analisi delle perdite per il macchinario TM4000, con i tempi relativi ad ogni causale.

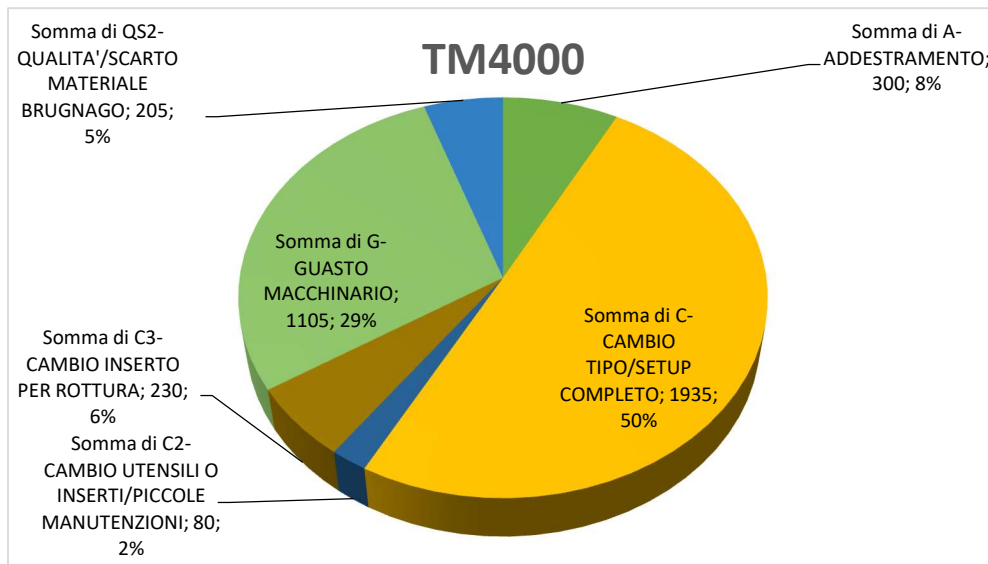


Figura 5 Analisi perdite del macchinario TM4000. Fonte: Brugnago S.R.L.

Dal momento che questo indice permette di identificare le cause che si celano dietro ai problemi, è possibile agire direttamente su quelle, al posto di correggere solo puntualmente il problema senza comprenderne la motivazione che lo genera.

Quantificabile attraverso il confronto con la situazione attuale relativamente a parametri come: spreco di inserti per manutenzione non accuratamente programmata, spreco di tempo per inefficienze nei fermi dei macchinari, inefficienza produttiva dovuta a basso effort degli operatori.

4.6 Soluzione 2: sviluppo di un modello di gestione dati e di un glossario di business

L'applicazione di soluzioni di data management all'interno dell'azienda Brugnago è stata possibile grazie alla collaborazione con Irion S.R.L. Un'azienda italiana che offre soluzioni di Data Management, attraverso la piattaforma all-in-one Irion EDM. La piattaforma e il suo paradigma dichiarativo rappresentano un salto generazionale rispetto ai sistemi tradizionali di data management. Permettono di trarre vantaggio dai dati e gestire, in modo agile e flessibile, l'evoluzione digitale del business. Il paradigma dichiarativo permette infatti di specificare alla piattaforma cosa debba essere fatto senza doversi preoccupare di come debba essere fatto. Tutte le attività che non richiedono un intervento umano vengono eseguite automaticamente dalla piattaforma sulla base della dichiarazione del risultato atteso. Un grande punto di forza della soluzione software offerta da Irion è la presenza di tutti gli strumenti necessari per il data management all'interno un'unica piattaforma, progettata coerentemente fin dall'inizio. In questo

modo funzionalità di data integration, data quality, data governance, lineage, data security, analytics, business intelligence e molto altro, sono tutte disponibili in un'unica soluzione.

4.6.1 Fasi di sviluppo del modello

Per lo sviluppo del modello attraverso la piattaforma di gestione dati Irion EDM, è necessario seguire una serie di passaggi fondamentali per la buona riuscita del progetto:

- Sviluppo del modello definendo elementi quali entità, relazioni e attributi al suo interno, in modo da creare un glossario di business per l'azienda. All'interno di un'azienda il linguaggio utilizzato è un tema fondamentale per gestire al meglio i dati e i processi. È particolarmente importante avere una chiara definizione dei dati, per questo motivo spesso viene sviluppato un vocabolario interno. Sviluppare e documentare standard sui dati, attraverso definizioni chiare e rigorose nella descrizione, riduce l'ambiguità e migliora le comunicazioni. Con la creazione di un modello è possibile anche lo sviluppo di un business glossary, che contiene le definizioni condivise di termini di business e le loro relazioni con altri dati. Nel caso di Brugnago, un glossario è importante per definire e descrivere chiaramente i dati e metadati di produzione, come macchinario, prodotto, pezzo, ordine e controlli di qualità come la rugosità. In generale la creazione di un business glossary permette di stabilire un linguaggio comune sui concetti base e sulla terminologia di business, ridurre i rischi legati ad una definizione non chiara di alcuni dati utilizzati in azienda, massimizzare la capacità di ricerca ed accesso ai dati e documenti aziendali.
- Integrazione dati nel modello. Nel modello sviluppato per Brugnago, per poter fare periodicamente analisi ed interrogazioni sui dati di produzione, devono essere integrati i dati con una periodicità fissa, in funzione di quello che l'azienda vuole monitorare e delle analisi che vuole svolgere. Questa operazione può essere svolta in più modi, tra cui:
 - Download dei dati attraverso file excel dai sistemi sorgente dell'azienda Brugnago per poi caricarli sulla piattaforma di Irion.
 - Creazione di un collegamento diretto attraverso Application Program Interfaces (API), che consente di trasferire i dati direttamente dal sistema sorgente alla piattaforma. In questo caso è necessario verificare che i protocolli del sistema sorgente permettano di creare il collegamento.
- Creazione di chiavi comuni. Un problema fondamentale da affrontare quando si parla di integrazione, riguarda le modalità con cui i dati provenienti da fonti diverse possano essere

messi in collegamento tra di loro. In questo caso la soluzione consiste nel trovare delle chiavi comuni tra le varie fonti, che permettano di mettere in comunicazione i dati durante il loro processo di integrazione e analisi. Ad esempio nel caso di Brugnago, considerando i dati relativi ad un ordine di produzione, le fonti sono collegate tra di loro attraverso il codice dell'ordine di produzione. Questo infatti è presente all'interno del sistema ERP in fase di programmazione degli ordini, durante le lavorazioni all'interno del software Overone e durante i controlli qualità. In questo modo durante tutto il ciclo di vita dei dati è presente una chiave di riferimento univoca che permette di mettere in comunicazione le fonti dati tra di loro.

- Implementazione del modello e definizione delle statistiche ed interrogazioni che vogliono essere poste allo strumento di analisi. Tra queste ad esempio:
 - Valutazione dei lotti prodotti nell'ultimo anno da uno specifico macchinario, considerando i controlli di qualità che sono stati fatti e prioritizzando gli errori che hanno portato più frequentemente alla non conformità di un pezzo.

4.6.2 Descrizione modello

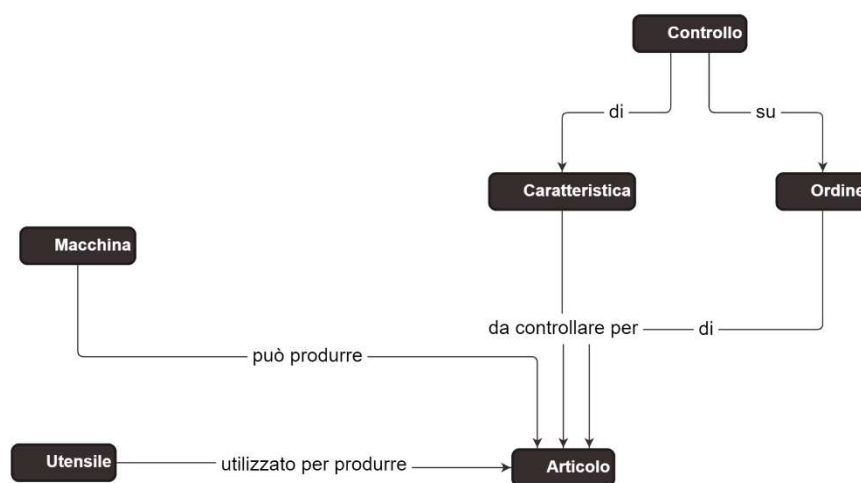


Figura 6 Modello per l'azienda Brugnago

Il modello sviluppato per Brugnago è mostrato in figura 6. Con esso si possono visualizzare in un grafo semantico le entità gestite nel perimetro della soluzione. Il modello definisce le entità rappresentate, come ad esempio articolo, ordine, caratteristica, e i loro mutui legami. È possibile comprendere attraverso una mappa di navigazione quali sono i prodotti completati da un determinato macchinario negli ultimi mesi. Partendo da una singola macchina si possono avere moltissime informazioni, tra cui: quali articoli può produrre, quali articoli sono stati prodotti, in

maniera tale da poter vedere, considerando la macchina, quale sia la situazione nel reparto. È possibile inoltre fare analisi relative ad uno specifico ordine che il sistema sta processando.

Il modello non viene sviluppato per avere una sua evoluzione in real-time, questo fattore dipende dalla frequenza di alimentazione dei dati. Il modello deve essere aggiornato periodicamente in momenti di consolidamento, come ad esempio un certo numero di volte al giorno. Lo strumento in questo modo fornisce un'istantanea in base ai dati che vengono forniti dalle diverse fonti nella situazione attuale. Ognuna di queste istantanee può essere salvata come playground, all'interno di una changeline. Quindi all'interno di un modello si potrà risalire ad una versione riferita ad una certa data e ora, permettendo così di analizzare il suo andamento e l'evoluzione nel tempo.

Il grande valore di questa soluzione di rappresentazione dei dati sta nel fatto che tutte le interrogazioni e analisi che un utente senza particolari competenze tecniche si vuole porre, possono essere visualizzate tramite mappe di navigazione generate dal modello. Ad esempio è possibile richiedere al modello che cosa sia stato fatto per questo determinato ordine. Nel momento in cui si apre la mappa di navigazione dell'ordine appaiono tutte le informazioni necessarie a riguardo. Oppure è possibile richiedere quali controlli di qualità siano stati fatti su un determinato ordine. Attraverso un'integrazione nel modello dei dati relativi ai controlli di qualità, viene generata una mappa di navigazione anche in questo caso. Si seleziona l'ordine, si sceglie la mappa di navigazione che ci interessa e si lavora su quella attraverso un determinato interrogativo.

L'immagine di seguito mostra, come descritto in precedenza, la rappresentazione in un grafo semantico di un'analisi relativa ad una specifica macchina denominata TM4000 e all'articolo martinetto con lo specifico codice.

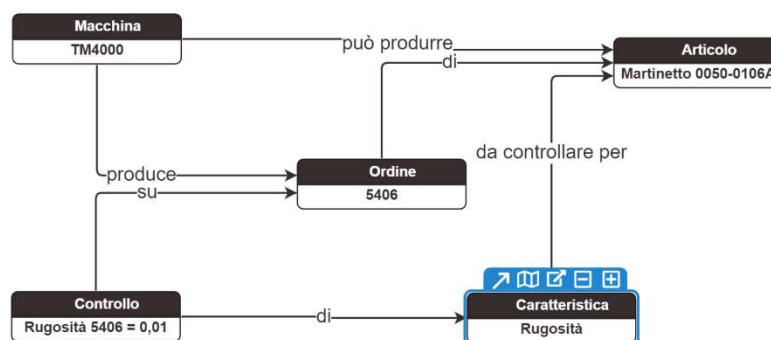
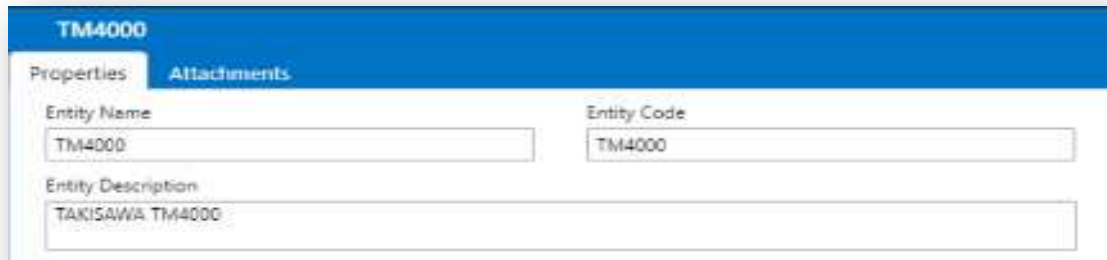


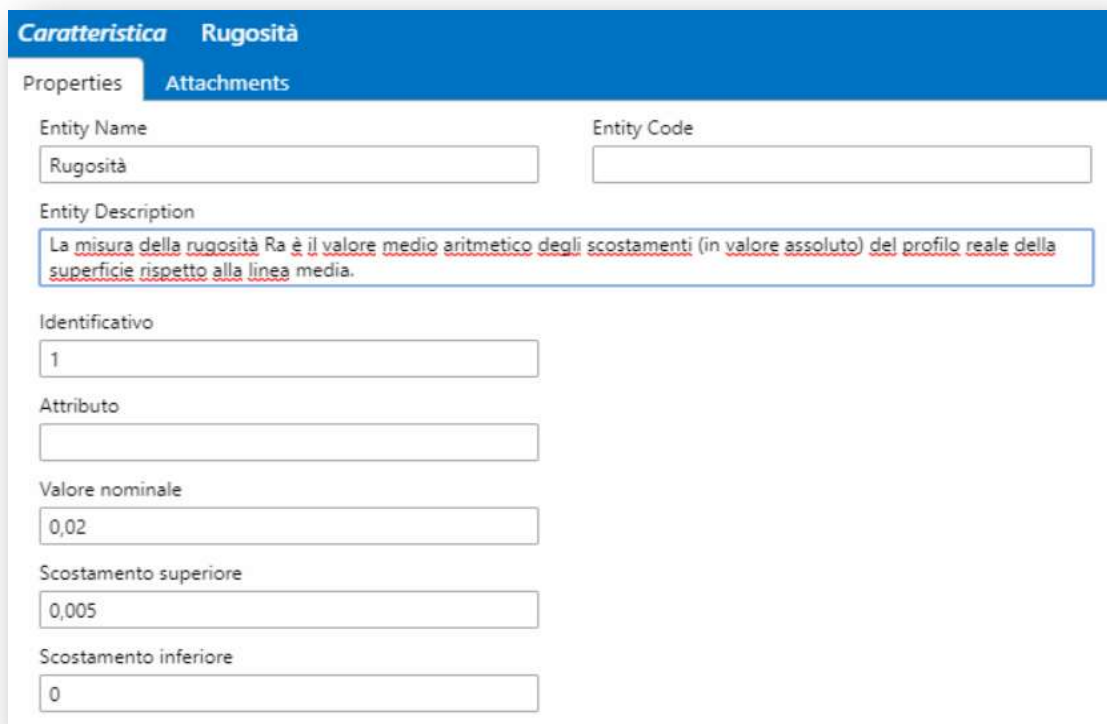
Figura 7 Grafo semantico di analisi del macchinario TM4000

Com'è possibile osservare dalle immagini sottostanti, cliccando su una entità come macchinario o caratteristica è possibile esplorare informazioni come il suo codice, la sua descrizione e i suoi attributi. Si valutano ad esempio per la caratteristica di rugosità gli scostamenti dal valore nominale, consentendo di identificare i pezzi che escono dalle tolleranze e per quale valore.



TM4000	
Properties	Attachments
Entity Name	Entity Code
TM4000	TM4000
Entity Description	
TAKISAWA TM4000	

Figura 8 Informazioni sul macchinario TM4000



Caratteristica Rugosità	
Properties	Attachments
Entity Name	Entity Code
Rugosità	
Entity Description	
La misura della rugosità Ra è il valore medio aritmetico degli scostamenti (in valore assoluto) del profilo reale della superficie rispetto alla linea media.	
Identificativo	
1	
Attributo	
Valore nominale	
0,02	
Scostamento superiore	
0,005	
Scostamento inferiore	
0	

Figura 9 Informazioni sulla caratteristica rugosità

4.6.3 Benefici

Qualificabili ma non quantificabili

- Possibilità di visualizzare e comprendere i dati in maniera più chiara ed immediata per utenti del Management, che probabilmente possiedono meno skills tecniche. Il management dell'azienda può in questo modo prendere decisioni in modo più consapevole ed avere una visione più chiara delle informazioni.
- Miglior catalogazione a sistema di macchinari, utensili, ordini, controlli di qualità a sistema grazie alla creazione di un glossario di business. Per l'azienda, conoscere gli utensili a sistema, rappresenta un'importante fonte di valore. Questa opportunità nasce dai dati che sono integrati nel modello dal gestionale e ad un chiaro e ben definito glossario. Ad esempio, conoscendo il codice del prodotto grezzo e il tempo di produzione, è possibile ricavare informazioni economiche utili sulla produzione. Per svolgere queste analisi sono necessarie una gestione del lessico accurata e l'introduzione di modelli di calcolo.
- Sviluppo di un sistema integrato a 360 gradi tra i vari reparti aziendali, soprattutto tra quelli che normalmente non scambiano informazioni tra di loro, conoscendo le definizioni degli utensili, degli articoli, dei macchinari, delle misure di qualità e molto altro.
- Possibilità di monitorare nel tempo lo sviluppo del modello e di costruire una base storica della situazione produttiva aziendale
- Flessibilità della soluzione:
 - Più opzioni di acquisto: On cloud e On Premises.

On cloud: con questa soluzione l'azienda paga un abbonamento, usufruendo della soluzione sotto forma di servizio. In questo modo non ci sono costi di investimento ma un costo per l'utilizzo della soluzione, che sarebbe imputato a costo operativo in conto economico per l'azienda, senza immobilizzazioni. La proprietà dei dati rimane dell'azienda e la sicurezza è elevata. Volendo l'azienda che fornisce la soluzione software può non conoscere i dati che vengono prodotti dal cliente.

On premises: in questo caso l'azienda acquista la licenza del software e lo installa su un server di sua proprietà. Dovendo gestire il software internamente si dovrà dotare di soluzioni hardware, come il server, con conseguenti investimenti iniziali più elevati. Oltre all'acquisto della macchina ci sono anche i costi di manutenzione della macchina e il costo delle risorse che dovrebbero seguire la soluzione software internamente all'azienda. L'aspetto

fondamentale riguarda l'immobilizzazione dell'investimento, a differenza della soluzione cloud, dove le spese rientrerebbero nel conto economico come costi operativi, in questo caso il software e il server verrebbero imputati come capitale immobilizzato.

Per un'azienda con le caratteristiche di Brugnago in termini di fatturato e numero di dipendenti, almeno all'inizio sarebbe più conveniente una soluzione di tipo cloud. Prima di tutto per i costi maggiori per l'acquisto del software e dei server necessari per il supporto, in secondo luogo attraverso una soluzione di questo tipo l'azienda potrebbe esplorare in maniera incrementale le possibilità offerte da uno strumento di questo tipo e valutare l'efficacia delle soluzioni e delle analisi che caratterizzano la piattaforma, e il rapporto tra il costo e il valore della soluzione introdotta in azienda. Il vantaggio di una soluzione on premises è rappresentato dalla proprietà dell'asset. Dopo un periodo di payback, il beneficio derivante dall'acquisto della soluzione sarebbe maggiore rispetto al noleggio del software. Il cloud, soprattutto quando non si parla di grandi organizzazioni risulta una soluzione molto utilizzata, dal momento che garantisce anche un elevato livello di sicurezza dei dati, appoggiandosi a sistemi come Amazon AWS, Microsoft Azure, Google.

- Possibilità di gestione della piattaforma da parte dell'azienda software limitando così l'effort necessario per lo sviluppo di competenze interne. In questo caso la soluzione viene definita Software as a service (SaaS), dove al cliente viene garantito tutto il servizio, inclusa la gestione del software. In questo modo il numero di risorse all'interno dell'azienda addette alla gestione si ridurrebbero. Inoltre soluzioni come quelle di Irion propongono servizi di application management, attraverso i quali il fornitore gestisce per conto dei clienti il caricamento e l'integrazione periodica da parte dei vari sistemi, monitorandone l'esito e in caso negativo eseguendo gli opportuni interventi di diagnosi e remediation del problema. Il cliente è supportato attraverso ticket per qualunque problema sorga relativamente all'utilizzo del software.

Quantificabili dopo l'adozione della soluzione

- Possibilità di creare statistiche, analisi ed interrogazioni in modo da supportare meglio le decisioni aziendali. Inoltre, attraverso l'utilizzo di semantic knowledge graph è possibile rappresentare graficamente le relazioni tra dati, informazioni e fonti dati.

Quantificabile attraverso la valutazione del numero di nuove analisi create e portate come materiale di supporto a riunioni manageriali/strategiche.

4.7 Le due soluzioni graficamente

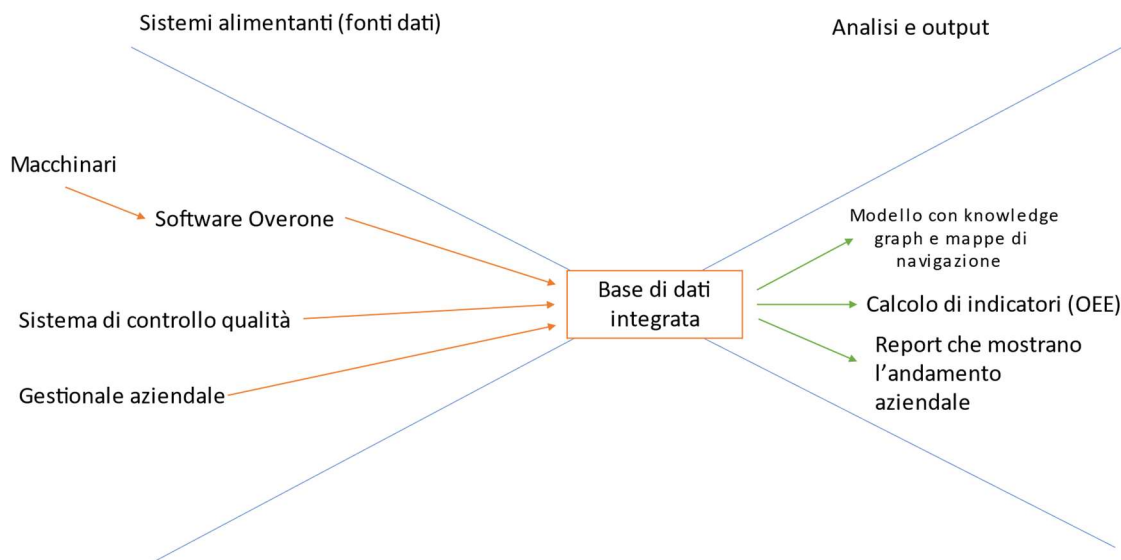


Figura 10 Grafico delle soluzioni proposte

Nella parte sinistra si possono riconoscere i sistemi alimentanti, ovvero le fonti di dati. Le frecce in colore arancione rappresentano fasi di integrazione dei dati. L'integrazione avviene infatti in due step principali.

1. Integrazione e standardizzazione dei dati provenienti dai diversi macchinari attraverso il software di raccolta dati Overone.
2. Integrazione dei dati in base alla tipologia di informazioni. Dal software Overone derivano i dati di produzione, dal gestionale dati commerciali, di pianificazione di ordini e produzione, dal macchinario di controllo qualità informazioni relative alla qualità dei dati.

Tutti i dati, se opportunamente integrati e messi in comunicazione tra di loro attraverso chiavi comuni di lettura (ad esempio il codice dell'ordine di produzione), rappresentano la base dati integrata. La prima soluzione si posiziona dal lato sinistro del grafico, facendo sì che l'integrazione dei dati svolta ora manualmente dagli operatori, sia invece svolta automaticamente dai sistemi, considerando il caso d'uso specifico dell'indice OEE.

La seconda soluzione, ovvero l'introduzione di una piattaforma di raccolta e gestione dati all'interno dell'azienda, permette di internalizzare la base dati integrata all'interno della piattaforma stessa. I dati opportunamente integrati e trasmessi alla piattaforma, saranno un patrimonio integrato di informazioni, da cui nascono i benefici, sul lato destro, derivanti dall'adozione di questa soluzione.

Tra questi:

- Modello di navigazione e rappresentazione grafica attraverso mappe di navigazione per mostrare i dati in maniera più semplice ed intuitiva.
- Calcolo di KPI tra cui si potrebbe inserire anche lo stesso OEE, anche su diversi livelli di analisi.
- Produzione di una serie di report che permettono di comprendere come stia andando l'azienda, rispetto a produzione, aspetti economico-finanziari, qualità dei prodotti, vendite.

In questo caso quindi, con l'adozione della soluzione software, la base dati e la raccolta dati potrebbe essere gestita completamente dal fornitore del servizio Irion, che con i suoi strumenti darebbe come output tutte le informazioni sulla destra del grafico.

4.8 Analisi dei costi

Per rendere possibili le soluzioni sopra descritte sono necessari alcuni investimenti per l'azienda manifatturiera:

- Introduzione di un software di monitoraggio della produzione e raccolta dati. Per valutare la fattibilità è stato considerato l'investimento fatto dall'azienda Brugnago per introdurre il monitoraggio con il software Overone su 22 macchinari di lavorazione. Le componenti principali dell'investimento sono relative a:
 - Tablet a bordo macchina
 - Lettori bar code
 - Schede wifi per la connessione
 - Costi di installazione ed addestramento

Totale di 33008 euro. Affinchè la gestione dei dati sia efficace, è necessaria una raccolta dati scrupolosa, che non può essere svolta manualmente dagli operatori nel reparto. Per questo motivo uno strumento di raccolta dati di produzione come il software Overone risulta un investimento fondamentale.

- Introduzione di una nuova architettura server. Brugnago ha dovuto sostenere l'investimento per una nuova architettura server. Il costo complessivo per l'acquisto e l'installazione dell'infrastruttura server è stato di 36700 euro. Questa risulta fondamentale non solo per le piattaforme di raccolta e gestione dei dati, ma anche per l'utilizzo di ERP aziendali. È importante una soluzione di questo tipo per la gestione a 360 gradi dell'infrastruttura IT per un'azienda che punta sulla digitalizzazione.

In alternativa è possibile valutare costi e benefici derivanti dall'utilizzo di una soluzione cloud come Amazon Web Services oppure Google Cloud, che permette all'azienda di doversi concentrare meno sull'infrastruttura IT e non dover installare un server fisico all'interno dell'azienda.

- La terza e importante decisione di investimento è legata al contratto software con un'azienda che offre servizi di data management, in questo caso Irion. Il costo annuo di noleggio della licenza dipende dalla soluzione, che può essere o meno in cloud. Nel caso di Brugnago è stata valutata più conveniente quella in cloud come soluzione iniziale. In generale il canone annuo di noleggio per la soluzione cloud parte da un minimo di 20000 euro. Tale costo cresce in funzione di alcuni parametri, per esempio
 - moduli sviluppati
 - potenza delle macchine server connesse
 - numero di utenti che hanno accesso al software

Il canone annuo di noleggio per un'azienda come Brugnago, con un numero contenuto di utenti connessi e dati condivisi, potrebbe quindi crescere di qualche decina di migliaia di euro, per arrivare ad un investimento annuo tra i 50000 e i 100000 euro. In ogni caso il valore della transazione dettagliato, è una variabile che dipende dalla contrattazione commerciale tra il cliente e l'azienda che offre il servizio software.

- Costo di gestione del cambiamento culturale aziendale, legato all'utilizzo dei dati nelle pratiche aziendali. Le voci di costo sono riconducibili a possibili corsi di formazione e certificazioni per i dipendenti coinvolti nella gestione dei dati. In particolare all'introduzione di una figura, già interna all'azienda o da inserire nell'organico, che sia il punto di riferimento per la gestione dei dati in azienda.

Quantificabile attraverso la valutazione di parametri come ad esempio: numero di dipendenti coinvolti nei corsi/certificazioni, costo del singolo corso di formazione, durata dei corsi.

4.9 Analisi del cambiamento culturale in azienda

La digitalizzazione e l'introduzione di procedure di data management all'interno dell'azienda, può comportare l'acquisizione di alcune risorse e competenze. Per quanto riguarda le competenze da sviluppare all'interno dell'azienda e quindi le possibili figure da inserire ci sono due modalità principali di affrontare l'accesso al mondo dei dati da parte dell'azienda. Una soluzione già citata in precedenza è rappresentata da un Software as a service (SaaS), dove al cliente viene garantito tutto il servizio, inclusa la gestione del software. In questo modo il numero di risorse all'interno dell'azienda addette alla gestione si ridurrebbero. La seconda alternativa è rappresentata dalla gestione in autonomia della soluzione, che comporta la presenza stabile all'interno dell'azienda di personale specializzato, con la conseguenza di dover sostenere un costo fisso, che deve essere giustificato dal suo pieno utilizzo.

All'interno di Brugnago attualmente i dati vengono monitorati dal manager di produzione, che non riesce però a dedicare il tempo necessario alle analisi. Questo è causato principalmente dalla mancanza di un team all'interno dell'azienda focalizzato sulla gestione ed analisi dei dati. Per questo la complessità e la completezza delle analisi è un punto aperto da sviluppare, dal momento che a valle delle analisi, vengono prese decisioni operative e strategiche. Manca quindi una figura che rappresenti il punto di riferimento per la gestione dei dati in azienda. Professionalità come Data Steward o Chief Data Officer (CDO) hanno un costo elevato, che potrebbe essere eccessivo in rapporto al fatturato dell'azienda. Queste figure solitamente svolgono soltanto una funzione strettamente legata alla gestione dei dati e per questo vengono introdotte all'interno organizzazioni strutturate come banche e assicurazioni, dove il core business è rappresentato dai dati stessi. Per un'azienda manifatturiera che sta iniziando ad introdurre al suo interno procedure di data management il rischio di non sfruttare appieno figure di questo tipo è elevato. La soluzione migliore è fare in modo che una figura già presente in azienda come ad esempio il manager di produzione o un'altra figura di middle management, attraverso un'adeguata formazione, possa incarnare in modo più responsabile il punto di riferimento per la gestione dei dati. In questo caso l'aspetto principale da affrontare è la formazione della figura, che possiede già competenze tecniche e manageriali, ma a cui mancano competenze di gestione dei dati. Un esempio di soluzione è rappresentato dalla gestione iniziale del sistema da parte del fornitore del servizio di data management, quindi in questo caso dall'azienda Irion. Con il passare del tempo l'azienda offre al cliente servizio di affiancamento

in modalità training on the job alla figura identificata per la gestione dei dati in azienda, che potrà così, continuando a svolgere i ruoli ricoperti normalmente, anche ricoprire la figura di Data Steward.

Per quanto riguarda le altre competenze IT e software da sviluppare all'interno dell'azienda, è interessante notare come Brugnago, tra i 14 dipendenti, non abbia identificato una figura con competenze di questo tipo. L'azienda, data la dimensione, il numero di dipendenti e il fatturato, ha deciso infatti di esternalizzare tutti i servizi di supporto presso i fornitori stessi delle piattaforme MES, del gestionale aziendale e dei server.

4.10 Benefici e creazione del valore per l'azienda

L'azienda Brugnago, grazie ai progetti di digitalizzazione sviluppati nel corso degli anni, ha a disposizione numerosi dati che possono essere utilizzabili per le analisi. L'implementazione di soluzioni di gestione, integrazione ed analisi dei dati creati da un'azienda manifatturiera, comporta investimenti. Di fronte all'introduzione di pratiche nuove come quelle proposte nella tesi, potrebbe quindi nascere scetticismo da parte degli operatori e del management. Per questo motivo è fondamentale evidenziare i benefici che derivano dalla realizzazione di questo progetto.

Benefici

Quantificabile dopo l'adozione della soluzione

- Creazione e utilizzo di dati di migliore qualità, con fonti affidabili e minori errori.
Quantificabile attraverso un confronto con la situazione attuale relativamente ad alcuni parametri come: numero di dati scorretti, numero di dati ridondanti, numero di dati ripetuti.
- Possibilità di effettuare analisi più accurate su inefficienze, sprechi e ritardi in produzione.
Quantificabile attraverso il confronto con la situazione attuale relativamente a parametri come: spreco di inserti per manutenzione non accuratamente programmata, spreco di tempo per inefficienze nei fermi dei macchinari, inefficienza produttiva dovuta a basso effort degli operatori.
- Migliore valutazione dei costi di produzione interni, che consente una valutazione del margine di vendita e del prezzo di contrattazione più precisa.
Quantificabile attraverso il confronto di fatturato, imputabile direttamente a prezzi di mercato più competitivi. Dati più affidabili permettono anche di monitorare e di valutare i costi di produzione con maggiore precisione

- Migliore valutazione degli impiegati, dei premi a produzione e delle azioni correttive.
Quantificabile attraverso un'analisi rispetto alla situazione attuale dei costi sostenuti dall'azienda per il personale e dal costo che può rappresentare per l'azienda il basso effort degli operatori.
- Possibilità di fornire al cliente, insieme al prodotto, informazioni su di esso legate alla tracciabilità logistica e qualitativa (controlli di qualità svolti sull'ordine, fasi di lavorazione ...). In questo modo viene instaurata con il cliente una maggiore fiducia, potendo quindi creare contratti più stabili e di lungo periodo. Questo implica vantaggio competitivo di fronte ai concorrenti sul mercato e possibilità quindi di acquisire nuovi clienti e quote di mercato.
Quantificabile attraverso una valutazione rispetto alla situazione attuale del numero di clienti per Brugnago e della durata media dei contratti di fornitura stipulati.
- Beneficio economico diretto derivante dalla vendita, insieme al prodotto, di dati e analisi dei dati.
Ad esempio per l'analisi predittiva e manutenzione preventiva vengono sviluppati dei modelli di machine learning da importanti organizzazioni software. Questi modelli richiedono una grandissima quantità di dati per essere allenati, che molto spesso chi li sviluppa non possiede. L'azienda Brugnago, archiviando dati di qualità sulla produzione, potrebbe pensare di vendere questi dati alle organizzazioni che sviluppano questi modelli.
Quantificabile attraverso un'analisi degli eventuali ricavi imputabili alla vendita diretta di dati da parte dell'azienda.

I benefici che sono stati indicati come attualmente non quantificabili sono specifici per la realtà aziendale in esame e non può essere fatto un confronto con una realtà simile a quella di Brugnago, dal momento che le soluzioni specifiche adottate per Brugnago non sono ancora state adottate da un'azienda con caratteristiche simili. La presenza all'interno dell'azienda di una base dati integrata con le giuste procedure, permette ora all'azienda di stimare dei target relativi ai singoli benefici, visto che sono disponibili dati di qualità elevata e fonti affidabili. Questi potranno poi essere ridimensionati a consuntivo in base al risultato ottenuto.

4.11 Condizioni necessarie per procedere con investimenti

Occorre ora analizzare quali sono le condizioni sotto cui ha senso, per aziende come Brugnago, procedere con investimenti. Il primo aspetto che consente al management di un'impresa di valutare le decisioni di investimento, riguarda la possibilità di risolvere i problemi interni, attraverso l'introduzione di nuove soluzioni tecnologiche. Compiendo un'analisi interna le imprese possono infatti comprendere come siano presenti inefficienze e possibili ambiti di miglioramento. Gli investimenti in digitalizzazione e data management risultano fondamentali qualora nell'azienda vengano riscontrati problemi di inefficienza o di mancanza di dati. Tali inefficienze possono essere dovute ai seguenti fatti: I dati all'interno dell'azienda spesso vengono registrati e trasferiti manualmente dagli operatori di reparto. Le analisi sono svolte utilizzando dati di bassa qualità che non provengono da fonti accurate, oppure alcune analisi non possono essere svolte per mancanza di dati, che potrebbero essere registrati grazie all'introduzione di nuovi macchinari più tecnologici. Dalle inefficienze interne nascono errori, che impediscono di svolgere correttamente analisi sui dati, o addirittura spesso non poterle fare. Questo in generale rallenta e diminuisce l'efficienza dei processi decisionali che si basano sull'utilizzo dei dati interni. Per questo motivo gli investimenti in digitalizzazione e Data Management risultano fondamentali per migliorare l'efficienza interna e la competitività sul mercato. Dalla presenza di ambiti di miglioramento dell'efficienza interna, nascono le condizioni necessarie per possibili decisioni di investimento, con il fine di migliorare le prestazioni interne, limitando le inefficienze e migliorando i processi.

Le aziende come Brugnago devono poi valutare i benefici e i possibili ritorni derivanti dagli investimenti effettuati. La presenza di effettivi benefici costituisce una condizione sufficiente per giustificare gli investimenti. I miglioramenti possono essere molteplici. Già soltanto la risoluzione dei problemi e inefficienze presenti in azienda, rappresenta un notevole beneficio. Per valutare se effettivamente venga compensato l'investimento fatto, è necessario svolgere un'analisi preventiva sulle risorse economiche perse a causa delle inefficienze presenti. Inoltre, gli investimenti in nuove tecnologie e risorse, permettono alle aziende come Brugnago di rinnovare il proprio business, concentrandosi maggiormente sui servizi offerti al cliente e sulla sua soddisfazione. Questo aspetto completa il ritorno sugli investimenti, che in questo caso sarebbero giustificati da contratti più stabili e duraturi nel tempo e da un possibile incremento di fatturato. In ogni caso i prodotti e servizi offerti dalle imprese sono molto diversi tra di loro, ma in tutte le aziende, le decisioni di investimento si basano sui possibili ritorni che ne deriverebbero. In sintesi è fondamentale quindi per le imprese

compiere un'analisi preventiva di benefici, possibili aumenti di fatturato e diminuzione dei costi, valutando se questi riescano a giustificare gli investimenti programmati.

Un aspetto da non tralasciare per imprese di queste dimensioni, che spesso hanno problemi di liquidità per investire, sono gli incentivi da parte dello Stato. In Italia, per esempio, per poter accompagnare le imprese durante il periodo di cambiamento tecnologico, è stato introdotto dal governo nel 2017 il Piano nazionale Industria 4.0. Industria 4.0 rappresenta una serie di misure, incentivi economici e agevolazioni sugli investimenti per lo sviluppo digitale delle imprese. La presenza di incentivi statali può infatti costituire un buon presupposto per investire in innovazione e digitalizzazione. Tali incentivi permettono di avere più facilmente accesso sia a risorse fisiche, come impianti e macchinari tecnologici, sia a nuove risorse umane e competenze, permettendo così una effettiva crescita dell'impresa.

Infine, potrebbero in futuro essere pubblicate normative per la gestione dei dati più stringenti di quelle attuali. Al momento infatti le normative più severe per la gestione dei dati, sono indirizzate a organizzazioni che fanno dei dati il proprio core business, come aziende nel settore finanziario, bancario e assicurativo. In vista di un possibile inasprimento delle leggi sulla gestione dei dati anche nelle aziende manifatturiere, nascerebbe un'ulteriore condizione che renderebbe necessario per le imprese investire in Data Management, per essere compliant con le normative.

Conclusioni

Con questo elaborato è stato possibile valutare, attraverso un caso concreto, i problemi e gli ostacoli che un'azienda manifatturiera deve superare per introdurre procedure di Data Management al proprio interno e i possibili benefici che ne derivano. Il lavoro svolto ha confermato le evidenze già mostrate nelle ricerche precedenti, ovvero che le aziende analizzate, nella loro situazione attuale, non sono pronte a questo cambiamento tecnologico. L'analisi della tesi ha dimostrato che i principali motivi sono riconducibili all'assenza di figure professionali e di competenze legate al Data Management nonché al mancato utilizzo di strumenti e di software di data analytics. In generale, quindi, mancano procedure consolidate di governance dei dati e anche una cultura aziendale focalizzata sul dato e sulla sua gestione. Il caso d'uso dell'azienda Brugnago ha permesso di comprendere come intervenire in pratica su questi problemi, proponendo due soluzioni diverse tra di loro ma indispensabili per una corretta gestione dei dati. Le due soluzioni, non solo forniscono esempi pratici di azioni da intraprendere all'interno dell'azienda per una migliore gestione dei dati, ma dimostrano anche un effettivo beneficio derivante da un loro corretto utilizzo. Nel primo caso la soluzione ipotizzata è stata l'introduzione di un'integrazione automatizzata di dati utili al calcolo dell'indice di produttività aziendale Overall Equipment Effectiveness (OEE). Questa prima soluzione, in generale, favorisce la creazione di un base dati integrata all'interno dell'azienda, che risulta necessaria per successive analisi. La seconda soluzione invece prevede un uso efficiente dei dati, correttamente integrati grazie alla creazione di un modello che utilizza una piattaforma di Data Management specializzata. L'utilizzo di questi strumenti dimostra un effettivo miglioramento nella visualizzazione, nella gestione e nell'analisi dei dati.

Il management di un'impresa deve avere un'immagine chiara degli obiettivi e una roadmap che definisce tutti gli step per raggiungerli, cercando di mantenere un livello elevato di qualità e accuratezza dei dati utilizzati nel processo decisionale. Attraverso un'analisi del volume e della tipologia di dati prodotti all'interno di un'azienda manifatturiera, il management pone l'attenzione su uno strumento che permetta di convertire dati provenienti da macchinari differenti, in modo da ottenere uno standard comune. Per poter introdurre soluzioni di Data Management all'interno dell'azienda è necessario effettuare investimenti. In particolare risulta fondamentale l'introduzione di strumenti che consentano la digitalizzazione e quindi una raccolta di dati precisa, come ad esempio il software di monitoraggio della produzione Overone analizzato nell'azienda Brugnago. È

necessario inoltre investire sul miglioramento dell'infrastruttura IT, attraverso l'acquisto di nuovi server e la configurazione corretta dei database, in modo da poter sostenere la digitalizzazione di impresa. In alternativa, si possono considerare costi e benefici derivanti dall'adozione di soluzioni IT in cloud al posto di ampliare le proprie soluzioni internamente. Dal lavoro di tesi è emersa, inoltre, l'importanza dell'accesso a servizi software di Data Management come quelli offerti dalla piattaforma dell'azienda Irion. Gli strumenti introdotti consentono al management di avere una visione più chiara dei dati, dei loro flussi e delle analisi ottenute, ma più in generale di trovare un vero loro utilizzo dal quale nasce la possibilità di creare valore per l'impresa. Ultimo, ma non meno importante, risulta lo sviluppo di competenze specifiche all'interno dell'azienda. È necessario, infatti, identificare figure specifiche che si occupino del coordinamento e della pianificazione di attività ed interventi da svolgere, affinché le procedure di Data Management funzionino in maniera corretta e sinergica. Professionalità di questo tipo come Data Steward o Chief Data Officer (CDO), tipiche delle grandi organizzazioni, possono tradursi in un costo elevato per un'azienda con caratteristiche di fatturato e dipendenti come Brugnago. Per un'azienda manifatturiera che sta iniziando ad introdurre al suo interno procedure di Data Management il rischio di non sfruttare appieno figure di questo tipo è elevato. La soluzione migliore è fare in modo che una figura già presente in azienda come ad esempio il manager di produzione o un'altra figura di middle management, attraverso un'adeguata formazione, possa incarnare in modo più responsabile il punto di riferimento per la gestione dei dati. Le figure aziendali che si occupano di gestione dei dati dovranno inoltre accompagnare l'introduzione della cultura del dato in azienda, altrimenti rischiano di essere isolate e di non riuscire a trasmettere completamente i valori e i vantaggi che l'utilizzo dei dati può portare con sé. Nello specifico, le comunicazioni da parte delle figure aziendali coinvolte devono mirare ad esprimere il valore tangibile e intangibile delle iniziative di Data Management e descrivere in che modo le sue funzionalità contribuiscano alla strategia e ai risultati aziendali. Soltanto in questo modo è possibile diffondere la cultura del dato in tutta l'azienda e a tutti i livelli, dal top management fino a livello operativo.

Per il management di impresa è quindi necessario evidenziare le condizioni sotto cui ha senso procedere con investimenti. Le nuove risorse e competenze introdotte in azienda dovrebbero consentire la risoluzione o quantomeno il miglioramento dei problemi presenti prima degli investimenti, facendo così nascere importanti benefici. Da parte di imprese con caratteristiche di dimensione e fatturato come Brugnago, la presenza di incentivi da parte dello stato, crea condizioni favorevoli per effettuare nuovi investimenti. Inoltre, un possibile inasprimento delle normative sulla

gestione dei dati aziendali, rappresenta una condizione necessaria di investimento in soluzioni di Data Management per le aziende nel settore manifatturiero.

In conclusione, per le imprese è necessario affrontare al meglio le sfide legate alla digitalizzazione, in modo tale da poter introdurre e consolidare competenze e strumenti al loro interno. Gli investimenti per aziende di questa dimensione e fatturato possono essere elevati, occorre quindi che il management svolga un'analisi interna sui costi e benefici derivanti dall'adozione di soluzioni in cloud per l'infrastruttura IT e dall'acquisizione di nuove competenze sotto forma di servizio esterno. In ogni caso queste risorse, se utilizzate nel modo migliore, permetteranno all'azienda di migliorare l'efficienza interna, il rapporto con i clienti ed il conseguente vantaggio competitivo sul mercato. I dati rappresentano un'importante opportunità per le aziende per innovare, ma è necessario concentrarsi sugli obiettivi di lungo periodo, cercando di rimanere aperti a nuove possibilità di business.

Fonti bibliografiche e sitografia

BIBLIOGRAFIA:

- DAMA – Dmbok data management body of knowledge 2nd edition (2020)
- Mauro Tuvo, Flavio Bordignon, Quaderni di management – Business Intelligence, E.G.V. edizioni, novembre – dicembre 2005, no. 18, p. 37-46
- Mariam Yasmin, Ekrem Tatoglu, Huseyin Selcuk Kilic, Selim Zaim, Dursun Delen, Big data analytics capabilities and firm performance: An integrated MCDM approach, Journal of Business Research, Vol. 114, 2020, p. 1-15
- Joseph Amankwah-Amoah, Samuel Adomako, Big data analytics and business failures in data-Rich environments: An organizing framework, Computers in Industry, Vol. 105, 2019, p. 204-212
- Zhou, Kaile, Fu, Chao, and Yang, Shanlin, Big Data Driven Smart Energy Management: From Big Data to Big Insights, Renewable & Sustainable Energy Reviews, Elsevier Ltd, Vol. 56, 2016, p. 215-225
- Sun, Yunchuan, Shi, Yufeng, and Zhang, Zhengjun, Finance Big Data: Management, Analysis, and Applications, International Journal of Electronic Commerce, Routledge Vol. 23, 2019, p. 9-11
- Anibal Reñones, Davide Dalle Carbonare, Sergio Gusmeroli, European Big Data Value Association Position Paper on the Smart Manufacturing Industry, European big data value association, 2018, p. 179-185
- Manlio Del Giudice, Veronica Scuotto, Armando Papa, Shlomo Y. Tarba, Stefano Bresciani, Merrill Warkentin, A Self-Tuning Model for Smart Manufacturing SMEs: Effects on Digital Innovation, Journal of Product Innovation Management, Vol. 38, 2020, p. 68-89
- Marco Pironti, Paola Pisano, Armando Papa, Technology Resilience and the S.T.O.R.M. Factory, SYMPHONYA Emerging Issues in Management, Vol. 2, 2018
- Fangchao Xu, Yongjian Li, Lipan Feng, The influence of big data system for used product management on manufacturing–remanufacturing operations, Journal of Cleaner Production, Vol. 209, 2019, p. 782-794

- Hannu Hannila, Joni Koskinen, Janne Harkonen, Harri Haapasalo, Product-level profitability: Current challenges and preconditions for data-driven, fact-based product portfolio management, Journal of Enterprise Information Management, 2019
- Jia Jia Liu*, Wen Hua Zhu, Eddie Soulier, Sebastien Remy, Research on Techniques of Intelligent Manufacturing Data Management, Advanced Materials Research, Vol. 1039, 2014, p. 449-455
- World Economic Forum, Data Excellence: Transforming manufacturing and supply systems, 2021

SITOGRAFIA

- <https://insidebigdata.com/2017/02/16/the-exponential-growth-of-data/>
- <https://www.networkworld.com/article/3325397/idc-expect-175-zettabytes-of-data-worldwide-by-2025.html>
- <https://www.idc.com/getdoc.jsp?containerId=prUS46286020>
- <https://www2.deloitte.com/us/en/pages/finance-transformation/articles/data-management-strategy.html#>
- <https://www.bloomberg.com/professional/blog/know-your-funds-data-and-risk-strategies-to-fuel-growth/>
- <https://www.datasciencedegreeprograms.net/industries/energy-management/>
- <https://www.scnsoft.com/blog/big-data-in-manufacturing-use-cases>
- <https://www.manufacturing.net/operations/article/13228439/using-big-data-analytics-to-improve-production>

Ringraziamenti

Mi è doveroso dedicare questo spazio dell'elaborato alle persone che hanno contribuito, con il loro instancabile supporto, alla realizzazione dello stesso.

Ringrazio infinitamente la mia famiglia che mi ha sempre sostenuto, appoggiando ogni mia decisione, fin dalla scelta del mio percorso di studi.

Ringrazio il mio relatore Emilio Paolucci, che in questi mesi di lavoro ha saputo guidarmi, con professionalità e suggerimenti pratici, nelle ricerche e nella stesura dell'elaborato.

Un ringraziamento speciale va al mio tutor dell'azienda Irion Mauro Tuvo, che mi ha aiutato a condurre le ricerche e sviluppare il caso di studio, sempre con estrema puntualità e disponibilità.

Ringrazio lo staff dell'azienda Brugnago S.R.L. per la disponibilità nel fornire informazioni per lo sviluppo della tesi.