

POLITECNICO DI TORINO

DIPARTIMENTO DI INGEGNERIA DELLA PRODUZIONE

Corso di Laurea Magistrale in Ingegneria Gestionale

Tesi di Laurea Magistrale

Tracciabilità e Data Silos: applicazione all'R&D farmaceutica



Relatore

Chiar.ma Prof.ssa Alessandra Colombelli

Candidato

Massimo Patti

A.A. 2020-2021

Sommario

Introduzione	5
Tracciabilità	7
L'importanza delle informazioni sui materiali.....	8
L'effetto sull'azienda della gestione delle informazioni sui materiali	9
Il costo del <i>Not Finding Information</i>	9
Scenario 1: <i>Time wasted searching</i>	9
Scenario 2: Costo della rielaborazione delle informazioni.....	10
Scenario 3: Costo Opportunità.....	11
Etichetta	12
Principali tipologie di etichettatura e confronti	12
Bar Code vs QR Code	14
RFID vs QR Code	14
Data Silos	16
Gruppo Chiesi Farmaceutici	18
La Storia.....	18
Mission	19
Le persone al centro.....	19
Lo spirito di squadra	19
Ogni operatore è Chiesi.....	20
Creazione di valore	20
Orientamento al cliente	20
Un'impresa affidabile	20
Sostenibilità	20
Società Benefit	21
Tracciabilità e Data Silos: il caso Chiesi Farmaceutici	22
Dipartimento CMC.....	23
Soluzione	24
BI Code	25
Metodologia di Approccio al Progetto	26
Bottom-up	26
POC e Prototipo.....	26
Metodo <i>Scrum</i> e <i>Agile Project Management</i>	27
Use Case	30
Soluzione	31
Web Application	31

Low-code/No-code Platform vantaggi e svantaggi.....	32
Cloud.....	34
BettyBlocks.....	35
Stampante.....	36
Fasi del processo.....	37
Fase 0.....	37
Fase 1.....	37
Fase 2- Creazione BI Code.....	38
Fase 2.1- Materiale scaricato da software aziendale.....	39
Fase 2.1.1.....	40
Fase 2.2- Materiale derivato da BI Code padre.....	41
Fase 2.2.1.....	42
Aggiornamento BI Code.....	43
Fase 3- Rintracciabilità.....	43
Analisi Economica.....	46
Costo di sviluppo.....	46
Riesame <i>Cost of not finding information</i>	48
Considerazioni.....	48
Conclusioni.....	49
Validazione e prospettive future.....	50
Appendice.....	51
Good Manufacturing Practice.....	51
Big Data.....	54
Le 3 V.....	56
La 4 ^a e la 5 ^a V.....	57
Come possono essere utilizzati i Big Data.....	57
I benefici aziendali e sociali dei Big Data.....	58
Data Integration.....	58
Il livello di integrazione dei dati: quattro approcci.....	58
Data Warehouse.....	59
Data Lake.....	59
Data Analytics.....	60
La rilevanza dei Big Data per i Manager e i CIO.....	61
Bibliografia e Sitografia.....	63
Ringraziamenti.....	65

Introduzione

Tra i problemi che si trovano ad affrontare le grandi imprese vi sono sicuramente quello della tracciabilità interna e quello della difficoltà nel far interagire tra loro diversi sistemi informativi aziendali, questo a causa della importante mole di oggetti/documenti da gestire e delle modalità attraverso le quali le informazioni vengono inglobate nell'azienda. Nasce quindi il problema del data silo, a cui segue quello della rintracciabilità, ovvero la capacità di ricostruire all'indietro l'intero processo di trasformazione di un prodotto incrociando dati e metadati provenienti da sistemi informativi non connessi tra loro.

Nel seguente elaborato andremo ad analizzare cosa si intende per tracciabilità e a osservare gli strumenti ausiliari per il raggiungimento dell'obiettivo. L'analisi farà riferimento ad aziende in cui possono sorgere problemi per il mancato tracciamento al loro interno, con un focus particolare su quelle farmaceutiche.

Successivamente andremo a studiare un problema che sta ottenendo sempre più rilevanza anche a causa dell'avvento dei Big Data ovvero quello dello *storage* dei dati e dei *data silos* che ne possono conseguire.

Nell'analisi presteremo particolare attenzione al *business problem* del Gruppo Chiesi Farmaceutici.

Dopo una prima descrizione dell'azienda si passa all'analisi della questione proposta, nel caso specifico, quello dei *data silos* e in particolare della difficoltà nella creazione di uno storico dati di tutte le entità (ad esempio campioni di materiali) che non sono normalmente registrate e tracciate sui sistemi di gestione informatica aziendali validati per uso farmaceutico (ovvero conforme alle norme GxP).

Nelle varie fasi del ciclo di vita, a uno stesso materiale viene associato un codice differente che può dipendere dal particolare stato di lavorazione, manipolazione o stoccaggio. Tuttavia, non essendo presente una chiave in grado di legare questi codici tra loro risulta difficile l'identificazione delle connessioni facendo così emergere la questione dei *silos*.

In risposta a ciò, è stata scelta una soluzione che possa essere attuabile cercando di ottimizzare i tempi e i costi sia nella fase di ricerca sia nel lungo periodo con l'effettiva applicazione della nuova metodologia. Si è dunque deciso di generare un codice alfanumerico denominato Business Intelligence (BI) code che permette di identificare in modo univoco tutto ciò che contiene informazioni che potrebbero essere soggette ad analisi in una qualsiasi fase del ciclo di vita di un qualsiasi manufatto. Questo codice fungerà da link trasversale per unire i dati di una stessa entità all'interno di silos (in questo caso i sistemi informativi) differenti.

Nel capitolo successivo viene esposta la metodologia d'approccio al progetto focalizzandosi soprattutto sul metodo agile, tipico dei progetti innovativi. Le scelte fatte durante questo lavoro hanno avuto come riferimento gli obiettivi futuri dell'impresa nell'ambito della ricerca e sviluppo, ovvero la propensione all'innovazione ed alla digitalizzazione dei processi aziendali.

Lo scenario di utilizzo analizzato riguarda la tracciabilità dei prelievi e dei test di laboratorio, nel quale non vi è un metodo specifico né uno standard di inserimento dati. Questo porta alla mancanza di uno storico completo e a un costo in termini di tempo necessario all'operatore di laboratorio per cercare informazioni, o colmare quelle mancanti, riguardo al materiale che sta analizzando.

Si è provato a trovare soluzione alla questione attraverso la creazione di una nuova metodologia di lavoro che non stravolge i processi tradizionali, ma li standardizza attraverso l'associazione del materiale al proprio "codice BI", al fine di creare una sorta di albero genealogico che permetta di tenere traccia dell'intero ciclo di vita. Questa nuova tecnica è stata supportata da un prototipo di applicazione, sviluppata attraverso una piattaforma low-code, della quale andremo ad analizzare vantaggi e svantaggi, che in caso di feedback positivo potrà essere sviluppato e ampliato ad altre aree aziendali.

Infine, verrà proposta un'analisi per lo più qualitativa dei costi e benefici dell'utilizzo di tale applicazione e con una prima individuazione di quelli che potrebbero essere i coefficienti di performance (KPI) da tenere sotto analisi durante la fase di test. Un'analisi quantitativa sarebbe risultata poco concreta non avendo un benchmark di riferimento con il quale fare le stime.

L'elaborato si conclude con un'ipotesi su quelli che potrebbero essere gli sviluppi futuri del progetto, per esempio, un'ulteriore automatizzazione e digitalizzazione dei processi ancora manuali e le possibili modifiche riguardanti le analisi economico-finanziarie che potrebbero essere implementate, in particolare con la revisione dei KPI definiti inizialmente e con la scelta di nuovi che possano portare all'identificazione a posteriori del modello benchmark.

Tracciabilità

La tecnologia ha già rivoluzionato il processo di produzione, portando a risparmi in termini di costi e risorse. Tuttavia, poiché il trend mondiale verso la digitalizzazione sta assumendo sempre più spessore, le organizzazioni possono iniziare a godere appieno degli ampi vantaggi della produzione potenziata dalla tecnologia attraverso solidi sistemi di tracciabilità.

Con il termine tracciabilità, “si intende la possibilità di risalire alla localizzazione, all’uso o alla storia di un prodotto attraverso delle identificazioni registrate” (ISO 8402). Tutte le fasi di produzione devono essere identificate e documentate, sia in relazione agli operatori di filiera che in relazione ai flussi materiali.

In un sistema produttivo, la tracciabilità assume un ruolo rilevante poiché in grado di assicurare: elevati standard qualitativi garantendo prodotti sicuri grazie ad un controllo qualità più efficace, la riduzione degli sprechi, un monitoraggio dettagliato di tutta la filiera produttiva e logistica in modo da rendere molto più facile identificare potenziali problemi, isolarne la causa principale e creare processi che possono essere ottimizzati nel tempo per un continuo miglioramento. È un fattore che sta rapidamente diventando una parte importante del processo di produzione per le aziende in diversi settori. Mentre alcuni credono che si applichi solo a prodotti che potrebbero essere soggetti a richiamo, come farmaci, cibo, automobili o aerei, la tracciabilità dovrebbe essere una parte dei processi di ogni produttore.

Tracciabilità (*tracking*) e rintracciabilità (*tracing*) si riferiscono a due concetti differenti ma strettamente interconnessi, il primo consiste nel processo attraverso il quale il prodotto viene seguito dall’inizio alla fine della filiera, così che a ogni fase vengano lasciate delle informazioni (tracce); il secondo, consiste nel processo opposto, attraverso il quale è possibile raccogliere le informazioni che sono state rilasciate in precedenza.

Per quanto riguarda le aziende farmaceutiche per assicurare la qualità del prodotto sono state inserite da anni una serie di regole, denominate Good Manufacturing Practices¹ (GMP), che descrivono i metodi, le attrezzature, i mezzi e la gestione delle produzioni per assicurarne gli standard di qualità appropriati.

Le GMP rappresentano quella parte dell’assicurazione di qualità che garantisce la consistenza tra i diversi lotti di produzione e la rispondenza ai parametri autorizzati nel dossier di registrazione.

Le verifiche ispettive presso le officine di produzione di medicinali sono condotte al fine di accertare che la produzione dei medicinali sia compiuta secondo criteri tecnici e metodologici tali da garantire la qualità del prodotto fabbricato.

Un mancato tracciamento del ciclo di vita dei documenti e dei materiali può essere ammissibile esclusivamente per tutte quelle entità che non richiedono per legge il rispetto delle norme GMP.

In particolare, per ciò che riguarda le analisi laboratorio in fase di sviluppo del prodotto non vi sono norme per la generazione di standard in grado di regolare i procedimenti di tracciabilità, nonostante questi standard siano molto importanti sia dal punto di vista delle coperture brevettuali sia da quello dell’ottimizzazione dei processi (know how).

Accade dunque spesso che i dati relativi a questi processi siano carenti di informazioni, che gli operatori impieghino molto tempo per raccoglierle, o addirittura, che siano stati smarriti.

¹ Appendice: *Good Manufacturing Practices*

L'importanza delle informazioni sui materiali

I dati sui materiali consentono di mettere in atto strategie migliori, con importanti implicazioni economico-finanziarie per le imprese in cui il costo dei materiali ammontano a decine/centinaia di milioni di euro (cifra che in aziende di progettazione può arrivare a miliardi di euro a livello globale).

Gli operatori possono risparmiare tempo, ridurre gli errori, assicurare decisioni basate su dati tracciabili e riutilizzare le conoscenze acquisite. In progettazione e analisi, gli ingegneri possono ottenere le informazioni esatte, in qualsiasi momento e in qualunque luogo, affinché possano essere prese decisioni più consapevoli e valutati in modo più preciso costi e rischi. Tutto questo porta a:

- prodotti migliori
- riduzione dei casi di insuccesso dei prodotti e del *time to market*
- massimizzazione dei benefici dell'uso degli strumenti operativi.

Se un'impresa fabbrica o progetta prodotti, allora utilizza dei materiali o le loro specifiche: le relative informazioni saranno necessarie in un determinato momento o addirittura durante tutto il ciclo di vita del prodotto stesso.



Figura 1: Esempi dell'uso di informazioni sui materiali nel ciclo di vita del prodotto

Secondo alcune interviste realizzate da Granta Design², molte organizzazioni devono ancora sviluppare un approccio aziendale sistematico nei confronti delle informazioni sui materiali. L'analisi che ha coinvolto ingegneri e manager del settore della progettazione ha individuato che più della metà degli intervistati aveva bisogno di gestire informazioni sui materiali, ma lavorava in organizzazioni che avevano limitate o assenti procedure per raggiungere questo obiettivo. Le

² <https://www.grantadesign.com/>

imprese citano di frequente problemi nel ricavare e riutilizzare i dati dei materiali anche all'interno di dipartimenti e team relativamente piccoli.

La capacità di acquisire informazioni o prendere una decisione e recuperare il suo intero contesto garantisce che i processi siano verificabili, contribuisce alla soddisfazione dei clienti e degli organismi di regolamentazione e rende possibile la rivisitazione e il completo utilizzo delle prove e delle analisi in un momento successivo.

L'effetto sull'azienda della gestione delle informazioni sui materiali

L'obiettivo delle migliori pratiche per la gestione delle informazioni sui materiali è quindi quello di avere fonti attendibili e completamente tracciabili per i dati aziendali sui materiali.

Il raggiungimento di questo obiettivo permette non solo di risparmiare tempo (gli operatori non devono più impiegare buona parte delle proprie ore lavorative a cercare i dati), ma anche di ridurre sprechi, errori ed evitare di ripetere analisi già svolte: "Fino all'80% del lavoro svolto in un dipartimento ingegneristico è uguale o molto analogo a quello svolto in precedenza"³.

Il costo del *Not Finding Information*

Quantificare il costo dell'impossibilità di trovare l'informazione da parte degli operatori è un'operazione alquanto difficile. Fattori che vanno dalla mancanza di formazione degli utenti, a sistemi aziendali incompatibili o a informazioni che semplicemente non esistono giocano un ruolo fondamentale; ecco perché il calcolo di questo costo sarà a sua volta costoso e spesso impreciso.

IDC ha sviluppato tre scenari che possono aiutare le aziende a stimare il costo di non trovare informazioni e il guadagno nella produttività che può essere raggiunto quando i *knowledge worker*⁴ dispongono di tool efficaci per la ricerca di informazioni.

Tutti questi scenari utilizzano ipotesi che devono essere adattate alle circostanze di ogni organizzazione.

Per lo sviluppo di questi casi, IDC assume uno stipendio annuo di \$ 80.000 per knowledge worker e una stima di 2,5 ore al giorno, o circa il 30% di giornata lavorativa, come tempo che impiega il lavoratore per cercare informazioni. Anche questo numero deve essere rettificato per riflettere le circostanze di ciascuna specifica impresa.

Ogni situazione mostra un potenziale motivo di fallimento di ricerca. È importante tener conto che questi scenari non possono lavorare insieme per ottenere una stima cumulativa. L'obiettivo di ognuno è quello di fornire una comprensione preliminare del problema.

Scenario 1: *Time wasted searching*

Non tutte le informazioni all'interno di un'azienda sono indicizzate o ricercabili. La ricerca preliminare suggerisce che circa il 35-50% delle informazioni disponibili all'interno di un'impresa non è indicizzato centralmente.

Assunzioni:

- Salario *knowledge worker* = \$ 80.000

³ Arthur D. Little

⁴ Professionista che si serve delle proprie conoscenze tecniche e specialistiche per risolvere i problemi di un'azienda.

- 1000 *knowledge worker* x 2.5 h/day in media per ricerca informazioni
- 40 h/week x 52 week = 2080 h/year

Calcolo dei costi:

$$\frac{80000 \frac{\$}{\text{year}} * \left(2.5 \frac{\text{h}}{\text{day}} * 5 \frac{\text{day}}{\text{week}} * 52 \frac{\text{week}}{\text{year}} * 50\% \right)}{2080 \frac{\text{h}}{\text{year}}} * 1000 \text{ worker} = 12'500'000 \frac{\$}{\text{year}}$$

Conclusion: un'impresa che impiega 1000 *knowledge worker* perde 24.000\$ a settimana, 12.5\$ milioni all'anno, a causa di un'incapacità di individuare e recuperare le informazioni

Scenario 2: Costo della rielaborazione delle informazioni

Per esporre il seguente caso IDC introduce il concetto di "*deficit di conoscenza*"⁵: ovvero una metrica che cattura i costi e le inefficienze derivanti principalmente dalla rielaborazione intellettuale, prestazioni scadenti e incapacità di trovare *knowledge resources*⁶. L'ampio studio di IDC ha permesso di stimare il costo medio di una gestione della conoscenza inefficace all'interno delle organizzazioni. Il deficit di conoscenza è stato tradotto in un costo medio di US \$ 5.000 annui per lavoratore. Dato in crescita rispetto alla data dello studio.

Uno studio⁷ di Kit Sims Taylor ha permesso di affermare che i *knowledge worker* spendono più tempo per ricreare inconsapevolmente la conoscenza esistente che per crearne di nuova (solo il 10% circa del tempo è speso nella creazione effettiva di nuova conoscenza).

Assunzioni:

- Salario del *knowledge worker* = \$ 80.000 di stipendio annuo
- 1000 lavoratori x \$ 5.000 all'anno*lavoratore (deficit di conoscenza)

Calcolo dei costi:

$$1000 \text{ worker} * 5000 \frac{\$}{\text{year*worker}} = 5'000'000 \frac{\$}{\text{year}}$$

Conclusion: un'impresa che impiega 1.000 *knowledge worker* spreca \$ 5 milioni all'anno perché i dipendenti spendono troppo tempo duplicando informazioni già esistenti nell'azienda.

⁵ European Knowledge Management Fact Book, IDC # 21511

⁶ informazioni o competenze all'interno di un'azienda che la rendono più preziosa o competitiva

⁷ Studio presentato alla International Conference on the Social Impact of Information Technologies, St. Louis, Missouri, <http://online.bcc.ctc.edu/econ/kst/BriefReign/BRwebversion.htm>

Scenario 3: Costo Opportunità

I primi due scenari si concentrano sui costi relativamente quantificabili per l'impresa del tempo sprecato a cercare e ricreare informazioni già esistenti ma che non è stato possibile trovare. Un altro costo tangibile ma piuttosto difficile da quantificare è il costo opportunità per l'organizzazione.

Assunzioni:

- Entrate per dipendente: \$ 500.000 all'anno o \$ 240 all'ora

Calcolo del costo opportunità:

$$50\% \text{ failed searches} * 240 \frac{\$}{h} * 2,5 \frac{h}{\text{day}} = 300 \frac{\$}{\text{day}}$$

Conclusion: il costo opportunità per l'impresa equivale a circa 300 dollari al giorno.

Altri costi imprevedibili per l'organizzazione includono il costo di decisioni sbagliate e delle vendite dovuto all'impossibilità di fornire informazioni ai clienti.

Mentre il costo di un cattivo processo decisionale sembra difficile da quantificare, non possiamo dire lo stesso sulle decisioni di progettazione ingegneristica, per esempio basate sull'utilizzo di componenti non più fabbricati che comportano una completa rielaborazione del prodotto.

Anche il tempo è un fattore da considerare. Se le informazioni non si trovano online, i dipendenti interrompono il lavoro dei colleghi per chiedere aiuto.

Tutti questi fattori sono difficili da quantificare, tuttavia devono essere considerati nella valutazione dei costi sostenuti dalla ricerca di informazioni inadeguate.

Sebbene i costi del *not finding informations* siano notevoli, sono nascosti all'interno dell'impresa e quindi raramente ne vengono percepiti gli effetti sui profitti. Tuttavia, le decisioni - se prese utilizzando informazioni scarse o errate - mettono in pericolo la vita dell'impresa.

Pertanto, è opportuno per l'impresa fornire i migliori strumenti di ricerca di informazioni disponibili e garantire che tutte le sue risorse intellettuali abbiano accesso ai dati indipendentemente da dove sono locati.

Etichetta

Ci sono due componenti principali che sono alla base di sistemi di tracciabilità efficaci. Questi sono:

- identificazione / codifica affidabile
- registrazione centralizzata. In realtà – soprattutto quando si parla di *big data* – la registrazione centralizzata potrebbe essere virtuale, o meglio, è sufficiente una gestione centralizzata di tutte le risorse che potrebbero essere fisicamente sparse ovunque, ma che vengono presentate come residenti in un'unica struttura centrale (virtuale).

Il processo richiede l'adozione di un sistema di identificazione o codifica che consenta l'etichettatura univoca di singoli prodotti, lotti o produzioni di una determinata area geografica. Ciò implica l'utilizzo di tecnologie come la marcatura permanente di parti o componenti nel momento in cui entrano nell'assemblaggio o direttamente dal fornitore. Le organizzazioni possono utilizzare, ad esempio, codici a barre 2-D, a matrice o tag RFID; a seconda della tecnologia utilizzata l'etichetta conterrà tipologie e quantità di informazioni diverse dalle altre.

L'etichetta contiene informazioni relative a una particolare fase di processo e di lavorazione del prodotto, tuttavia, anche nel caso in cui sia ricca di informazioni, non ha valore se considerata isolatamente; deve essere letta e decodificata attraverso sistemi di lettura integrativi per estrapolare i dati e contesto. Grazie a questi riusciamo a legare le informazioni derivanti dalle varie fasi di lavorazione e a riferirsi a un particolare step del ciclo di lavorazione nel momento in cui vi è necessità.

Nel paragrafo successivo verranno espone e confrontate le tre principali tecnologie applicabili sull'etichetta citate sopra.

Principali tipologie di etichettatura e confronti

➤ **Codice a barre lineare:** Il codice a barre è un codice di identificazione costituito da un insieme di elementi grafici a contrasto elevato. Tra i tipi più diffusi in Italia, vi è il codice EAN (*European Article Number*) utilizzato nella grande distribuzione, seguito dal Farmacode o codice 32 adottato per l'identificazione dei farmaci e delle specialità vendibili al banco nelle farmacie. Altri codici legati alla grande distribuzione sono lo Universal Product Code: utilizzato per i prodotti importati da Regno Unito, Australia, Nuova Zelanda, Canada, Stati Uniti; e i codici a barre GS1 DataBar, utilizzabili dal 2014, caratterizzati dalla capacità di memorizzare un maggior numero di informazioni in minor spazio.

Nel settore industriale hanno trovato grande diffusione il codice 128, il codice 39 e il 2/5 (si legge 2 di 5) interlacciato. A seconda del tipo di codice adottato vi sono dei limiti nel ridimensionamento, nel numero e nel tipo caratteri rappresentabili.

Il codice a barre lineare è composto da un susseguirsi di barre e spazi. La barra è l'elemento scuro del codice a barre, formato da un segmento verticale; lo spazio è l'elemento che separa due barre tra loro.

Per modulo si intende la larghezza dell'elemento (barra o spazio) più stretto. Gli altri spessori sono un multiplo del modulo.

Ciascuna sequenza di barre e linee deve essere preceduta e seguita da un'area di riposo (*quiet zone*) ovvero un'area chiara. Per permettere di capire se il codice a barre è stato letto per intero, l'informazione è preceduta da un carattere di start e seguita da un carattere di stop. Alcuni codici a barre hanno i caratteri di start e di stop uguali tra loro (es. CODE 39), in altri sono diversi (es.

code 128). La caratteristica di avere i caratteri di start e stop diversi è usata da alcuni codici a barre per determinare la direzione del codice a barre e quindi permettere la lettura da entrambi i sensi. Per permettere la correttezza della lettura del codice a barre, è previsto un codice di controllo. In alcuni codici a barre il codice di controllo è esplicito e fa parte dell'informazione (ES, EAN e UPC), in altri, il codice di controllo è implicito e non fa parte dell'informazione (es. Code 128).

Si definiscono codici a barre discreti, i codici a barre lineari in cui solo le barre portano l'informazione. Nei codici continui sia le barre che gli spazi portano informazioni. Nei codici a barre discreti, ogni carattere del codice a barre ha lo stesso numero di barre e spazi e ne rende quindi facile la rappresentazione con un font di caratteri (es. Code 39).

- **QR code o codice a barre a matrice:** è un codice a barre a due dimensioni, composto da moduli neri disposti all'interno di uno schema bianco di forma quadrata, impiegato in genere per memorizzare informazioni destinate a essere lette tramite apparecchi fotografici e smartphone. In un solo crittogramma possono essere contenuti fino a 7.089 caratteri numerici o 4.296 alfanumerici. Genericamente il formato matriciale è di 29x29 quadratini e contiene 48 alfanumerici. Il nome "QR" è l'abbreviazione dell'inglese Quick Response ("risposta rapida"), in virtù del fatto che il codice fu sviluppato per permettere una rapida decodifica del suo contenuto.

I codici QR possono contenere sia indirizzi Internet, che testi, numeri di telefono, o SMS. Sono leggibili da qualsiasi telefono cellulare e smartphone munito di un apposito programma di lettura (*QR reader*).

Denso Wave, la compagnia giapponese che ha sviluppato questo tipo di codice, ha reso pubblico l'uso della tecnologia QR con licenza libera, su Internet è possibile trovare programmi gratuiti sia per la lettura (decodifica) che per la scrittura (codifica) dei codici QR.

Dalla fine degli anni 2000 i programmi di lettura dei codici QR sono spesso già integrati nei telefonini dai relativi produttori.

Esistono comunque molti siti web che offrono gratuitamente lettori per cellulari.

Per leggere un codice QR è sufficiente inquadrarlo con la fotocamera del cellulare dopo aver aperto il lettore. Per quel che riguarda la scrittura, esistono diversi siti che consentono la libera produzione di codici.

Nei codici QR è applicato il codice Reed-Solomon per la rilevazione e correzione d'errore: nel caso in cui il QR fosse in parte danneggiato, per esempio da macchie o graffi sul supporto cartaceo, l'applicazione Reed-Solomon permette di ricostruire i dati persi, ripristinando, durante la decodifica, fino al 30% delle informazioni codificate.

- **Radio-Frequency IDentification (RFID):** si intende una tecnologia per l'identificazione e/o memorizzazione automatica di informazioni inerenti a oggetti, animali o basata sulla capacità di particolari etichette, chiamate tag, di memorizzare dati, e sulla capacità di queste di rispondere all'interrogazione a distanza da parte di appositi apparati chiamati reader.

L'identificazione avviene attraverso radiofrequenza, grazie alla quale un reader può comunicare e/o aggiornare le informazioni contenute nei tag che sta interrogando; infatti, nonostante il nome, un reader non è solo in grado di leggere, ma anche di scrivere informazioni. In un certo senso, i dispositivi RFID possono essere quindi assimilabili a sistemi di lettura e/o scrittura senza fili con svariate applicazioni.

Bar Code vs QR Code

Sebbene sia i codici QR che i codici a barre abbiano lo scopo fondamentale di memorizzare le informazioni relative a un determinato prodotto o organizzazione, ci sono alcune differenze importanti tra loro.

- ❖ Un'azienda può memorizzare molto di più in un unico codice QR che in un codice a barre; quest'ultimo contiene informazioni solo in direzione orizzontale, un QR può farlo anche verticalmente (questo è il motivo per cui i codici QR sono indicati come bidimensionali). Nello specifico: un codice QR può contenere fino a cento volte la quantità di informazioni di cui è capace un codice a barre.
Un codice a barre standard potrebbe essere in grado di indicare un numero di prodotto e la sua posizione prevista. Un codice QR, d'altra parte, sarebbe in grado di fornire le condizioni del prodotto, quando è stato realizzato, eventuali riparazioni effettuate e molto. Ciò significa che un'azienda può ottenere una comprensione più precisa del proprio inventario in generale.
- ❖ Oltre a trasportare più informazioni, il codice QR richiede meno spazio rispetto al *barcode*: un codice a barre convenzionale può richiedere fino a dieci volte la quantità di spazio di stampa di un codice QR contenente la stessa quantità di informazioni.
- ❖ Un codice QR è in grado di essere letto a 360 gradi, da qualsiasi direzione, eliminando così ogni interferenza ed effetti negativi dagli sfondi
- ❖ Minor margine di errore durante la scansione di articoli o codici sporchi o danneggiati da parte del QR code.



Figura 2: Punti di forza e debolezza del QR

RFID vs QR Code

Dal paragrafo precedente è emersa la maggior efficienza di una tipologia di tecnologia rispetto all'altra.

Andando ora a confrontare i QR con i tag RFID, non si osserva, al contrario del caso precedente, la definizione di un codice prevalente; l'azienda sceglierà la tecnologia in base all'utilizzo che ne deve fare.

- ❖ L'etichetta RFID permette l'accesso alle informazioni senza necessitare di una connessione Internet.
- ❖ I tag RFID possono contenere una quantità superiore di informazioni rispetto alle altre tecnologie.
- ❖ Poiché i tag e gli scanner RFID utilizzano la radiofrequenza, possono potenzialmente far risparmiare agli imprenditori molto tempo e denaro.

- ❖ Gli scanner RFID possono leggere più tag contemporaneamente, anche da una certa distanza e anche quando il tag è oscurato da un muro o da una scatola. Puoi anche scansionarli da qualsiasi direzione, in modo che i dipendenti del magazzino non debbano perdere tempo aggiuntivo per sistemare ogni scatola nel lato corretto per la scansione.
- ❖ I tag RFID sono più costosi dei codici a barre o dei codici QR. Inoltre, poiché la maggior parte dei telefoni cellulari non è dotata di funzionalità NFC, è necessario investire denaro extra in scanner RFID, senza utilizzare gli smartphone dei dipendenti.
- ❖ I tag RFID possono anche essere “imprevedibili”. Poiché contengono microchip, sono vulnerabili all'hacking da parte di malintenzionati che desiderano rubare le informazioni sui prodotti.
- ❖ I tag RFID possono anche essere inaffidabili se utilizzati in prossimità di un magnete, liquido o determinati tipi di metallo, quindi non saranno un'ottima scelta per tutti i tipi di prodotti.

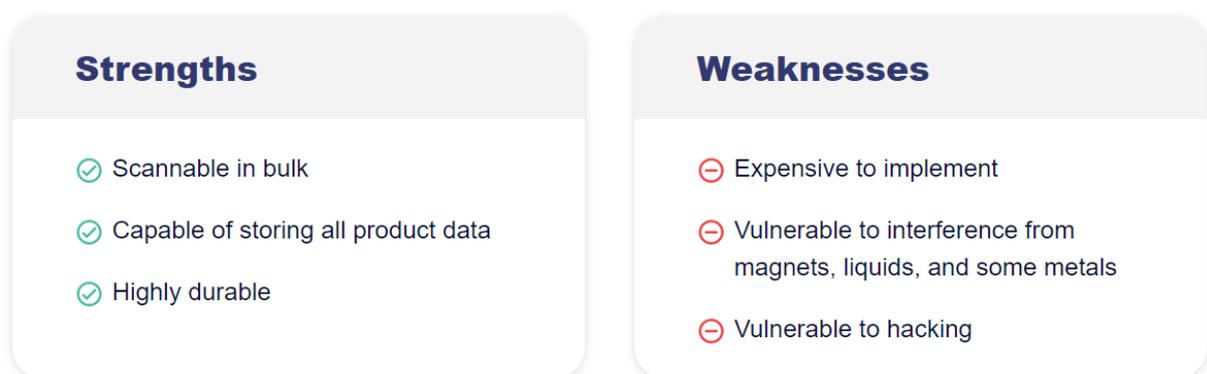


Figura 3: Punti di forza e debolezza dell'RFID

Data Silos

Più l'azienda è grande e più funzioni aziendali vengono ricoperte al suo interno, maggiore sarà la necessità che le diverse aree aziendali riescano a comunicare agilmente tra di loro in modo da avere un facile e reciproco accesso ai dati.

Tuttavia, questa non è una situazione così semplice da ottenere a causa di limiti aziendali che portano alla creazione dei cosiddetti "*Data Silos*", ovvero archivi di dati che restano sotto il controllo del reparto aziendale che li ha generati e sono isolati dal resto dell'organizzazione.

Il concetto di *silo* è associato a qualsiasi sistema non in grado di interagire con altri della stessa azienda a causa di una scarsa, o addirittura nulla, integrazione che porta alla creazione di ambienti di gestione dati isolati.

I *silos* rappresentano l'approccio tradizionale allo storage dei dati e tendono ad emergere in modo spontaneo, all'interno dell'azienda, nelle organizzazioni dove ogni linea di business raccoglie i dati in base agli utilizzi che ne deve fare e agli obiettivi che deve perseguire. L'archiviazione delle informazioni viene suddivisa per area aziendale, i dipartimenti sono isolati tra loro, non vi è comunicazione e di conseguenza i dati non sono integrati. Solo alcuni gruppi di persone all'interno dell'organizzazione possono accedere a determinati dati o alle relative origini, mentre le altre non sono in grado di ottenere informazioni sufficienti sul contesto dei dati o sui relativi metadati, incluso un intero gruppo di attributi come discendenza, significato, glossario aziendale, modifiche apportate o origine.

Le principali cause identificate sono le seguenti

1. Le Persone. Quando si tratta del personale, la cultura e le strutture organizzative sono le sfide più rilevanti. Spesso, i reparti sono isolati e lontani gli uni dagli altri, soprattutto nelle imprese più grandi.

E mentre questo è l'ordine naturale nella maggior parte dei casi, e anche visto come necessario per il funzionamento delle aziende, il problema è che le informazioni finiscono per essere legate a un particolare dipartimento o team e non fluiscono naturalmente da un reparto all'altro.

2. La Tecnologia. Le diverse tecnologie non parlano tra di loro. Potrebbero, ma sono condizionate dalle persone e dai processi a cui sono legate.

I *silos* di dati non sono un problema nuovo, nel 2012, McKinsey ⁸affermò che "i dipendenti trascorrono 1.8 ore ogni giorno, in media 9.3 ore a settimana, a cercare e raccogliere informazioni".

Tuttavia, oggi il problema è diventato rilevante in seguito all'emergere del fenomeno dei *Big Data*. Se nei decenni scorsi poteva essere considerato normale creare dei *silos* di dati, separati per ogni funzione aziendale, oggi, con l'avvento dei Big Data⁹, è necessario superare questa configurazione architettuale.

⁸ McKinsey & Company è una multinazionale di consulenza strategica

⁹ Approfondimento in Appendice: *Big Data*

La presenza dei *silos* viene riconosciuta come un ostacolo per le attività di business. Le aziende devono cercare quindi una strategia che permetta di abbatterli e che miri alla collaborazione, all'accessibilità e all'efficienza.

Gruppo Chiesi Farmaceutici

La Storia

L'azienda nasce a Parma nel 1935 con l'acquisto del Laboratorio Farmaceutico Parmense da parte del farmacista Giacomo Chiesi.

All'inizio degli anni Quaranta il laboratorio vede un primo ampliamento con le prime esportazioni di farmaci in Eritrea e in Austria

Nel 1944 l'attività viene fermata dall'esplosione di una bomba che distrugge buona parte dell'edificio.

Al termine della guerra l'attività riprende, viene assunto il primo collaboratore scientifico e viene lanciata la prima pomata a base di penicillina.

Nel 1955 viene inaugurata la prima fabbrica di Chiesi, conta 50 dipendenti e una produzione allargata di medicinali di successo.

Nel 1966 passano a capo dell'azienda Alberto e Paolo Chiesi, figli di Giacomo, i quali iniziano un processo di espansione ed internazionalizzazione che si concretizza nell'apertura in Brasile della prima sede estera alla fine degli anni Settanta.

Nel 1979 viene lanciato il Clenil, un prodotto cortisonico per l'asma, le riniti allergiche e le infiammazioni delle vie respiratorie in genere. Il farmaco riscuote grande successo, grazie soprattutto all'inefficacia della maggior parte dei farmaci dei competitor per questo tipo di patologie; nasce così l'impegno di Chiesi Farmaceutici nell'area delle malattie respiratorie.

nel 1989 il settore terapeutico d'interesse si estende anche all'area degli antiinfiammatori e antidolorifici tanto che nasce Brexin, un antiinfiammatorio dall'azione rapida e con buona tollerabilità gastrica che in poco tempo viene commercializzato in oltre 60 paesi, ottenendo inoltre la certificazione per la commercializzazione negli Stati Uniti e un brevetto internazionale.

Inizia così la nuova spinta globale dell'azienda, vi è una forte internazionalizzazione, sia in ambito europeo che extraeuropeo. Dal 1987 ad oggi Chiesi approda, in ordine temporale, in Pakistan, Francia, Spagna, Grecia, Gran Bretagna, Austria e paesi dell'Est europeo, Stati Uniti, Germania, paesi del Magreb, Turchia, Bulgaria, in Cina e nei paesi Scandinavi. In questi anni i prodotti storici, nelle aree terapeutiche respiratorie e cardiovascolari, continuano la loro affermazione, ma vengono lanciati anche nuovi prodotti per le patologie neonatali, muscolo-scheletriche e vascolari.

Con l'inizio del terzo millennio entra in azienda la terza generazione Chiesi, i figli di Alberto e Paolo: Alessandro, Andrea, Giacomo e Maria Paola. Con il loro ingresso si aprono ulteriori nuove linee di ricerca e sviluppo: la tecnologia formulativa Modulite® per i prodotti respiratori in spray eco-compatibili, e in particolare Foster®, combinazione antiasma in bomboletta spray con nuovi propellenti, nonché lo sviluppo dell'inalatore per polveri NEXTHaler®, considerato il più innovativo della categoria. Nel 2008 con la fondazione di Holostem Terapie Avanzate srl, Chiesi entra come pioniere nel mondo della medicina rigenerativa e nel 2013, con l'acquisizione della danese Zymenex, entra in quello delle biotecnologie, posizionandosi oggi all'apice dell'innovazione nel settore bio-farmaceutico.

Mission

“L’azienda vuole essere riconosciuta come Gruppo internazionale incentrato sulla ricerca, in grado di sviluppare e commercializzare soluzioni terapeutiche innovative che migliorino la qualità della vita delle persone. Intendiamo mantenere un team imprenditoriale di elevata qualità, caratterizzato da spirito di collaborazione e sicurezza in sé stesso. Il nostro obiettivo è unire impegno verso i risultati e integrità, operando con responsabilità sia dal punto di vista sociale, sia da quello ambientale.”



Figura 4: Raffigurazione dei principali valori dell'impresa

Le persone al centro

L’azienda è convinta che le persone siano il fattore più importante e per questo vuole promuovere e soddisfare nel tempo il bisogno di sviluppo, prestando sempre attenzione alle qualità dell’ambiente di lavoro e al benessere collettivo.

Lo spirito di squadra

Per il raggiungimento dei propri obiettivi l’azienda ritiene che le singole persone da sole non bastino, ma che il successo passi attraverso l’orientamento e il contributo individuale all’evoluzione di un team coeso e integrato, con la consapevolezza che il successo del team sia più importante del successo del singolo.

Ogni operatore è Chiesi

Ogni operatore è parte integrante dell'impresa, artefice del suo futuro e testimone dei valori che la distinguono. Ognuno è responsabile del sostegno della reputazione dell'azienda in ogni contesto e con ogni interlocutore.

Creazione di valore

Le leve su cui fa affidamento Chiesi per creare valore nel presente e nel futuro sono l'imprenditorialità e l'innovazione. L'imprenditorialità intesa come assunzione responsabile di rischio nel cogliere le opportunità coerenti con i principi e l'attivazione consapevole delle proprie competenze e risorse per sfidare il presente. L'innovazione intesa come capacità di cambiare concretamente le cose attraverso un processo costante di valorizzazione delle intuizioni e delle idee della nostra comunità professionale.

Orientamento al cliente

Chiesi vuole essere costantemente attenta, sensibile e dedita ai bisogni dei suoi clienti generando la loro più ampia soddisfazione, nel tempo più breve di risposta, attraverso l'eccellenza dei suoi prodotti e servizi e la passione nelle relazioni che sa quotidianamente offrire.

Un'impresa affidabile

Vuole essere riconosciuti da tutti i nostri interlocutori come un'impresa che promuove e adotta a tutti i livelli comportamenti trasparenti e coerenti con le norme, le regole e che si ispira ai principi della responsabilità sociale.

Sostenibilità

Nell'azienda trova grande rilevanza anche il tema della sostenibilità. Chiesi è il più grande gruppo farmaceutico internazionale a ricevere la certificazione di B Corp¹⁰, riconoscimento di alti standard sociali e ambientali.

I temi della sostenibilità sono oggi fondamentali nella visione strategica dell'Azienda, che si impegna a:

- adottare un sistema di governance basato sulla trasparenza e responsabilità;
- favorire lo sviluppo professionale e personale delle persone ed il loro benessere;
- proteggere il pianeta e minimizzare gli impatti prodotti dalle proprie attività;
- rispondere in maniera efficace ai bisogni medici non soddisfatti dei pazienti;
- servire la comunità e le persone in maggior stato di bisogno.

¹⁰ Aziende che perseguono sia il profitto che un beneficio comune, secondo i più alti standard di performance sociale e ambientale, trasparenza pubblica e responsabilità. Per diventare una B Corp bisogna ottenere un punteggio minimo nel B Corp Assessment, una valutazione rigorosa dell'impatto che un'azienda ha sui suoi lavoratori, clienti, sulle comunità e sull'ambiente.

Per questi obiettivi è stato creato il programma “We ACT - We Actively Care for Tomorrow”, che esprime la volontà di prendersi cura attivamente della società e dell'ambiente.

Questo programma di sostenibilità crea continuità con i valori e la missione aziendale di unire impegno verso i risultati e integrità, operando responsabilmente sia dal punto di vista sociale che ambientale.

Società Benefit

Nel 2018, Chiesi diventa una Società Benefit secondo la legge italiana n. 208/2015, integrando nel proprio statuto 4 finalità specifiche dedicate ai propri pazienti, alla comunità, ai dipendenti, ai fornitori e all'ambiente in cui opera, oltre ad una finalità generica di beneficio comune.

Le Società Benefit rappresentano uno dei modelli più avanzati in termini di impatto positivo sociale e ambientale d'impresa. Sono società for profit che esplicitano all'oggetto sociale un duplice obiettivo: creare valore sia per gli shareholder che per gli stakeholder. Insieme costituiscono un movimento globale di 'Purpose Driven Businesses' che ha l'obiettivo di diffondere un paradigma economico più evoluto, che vede le aziende come protagoniste nel rigenerare la società e la biosfera. La visione delle Società Benefit è di innescare una competizione positiva tra tutte le aziende, perché siano misurate e valutate nel loro operato secondo uno stesso metro: l'impatto positivo sulla società e il pianeta, oltre ai risultati economici.

Le Società Benefit devono presentare una Relazione di Impatto in cui indicare ogni anno degli obiettivi rispetto alle finalità specifiche poste all'interno del proprio statuto.

Tracciabilità e Data Silos: il caso Chiesi Farmaceutici

Il problema evidenziato dall'azienda, in particolare dal reparto CMC (Chemistry Manufacturing & Controls), è quello della difficoltà nel tracciare e seguire il "ciclo di vita" di tutte quelle entità fortemente eterogenee, come materiali e documenti (ad esempio: campioni di laboratorio con relative analisi), per il quale non era richiesta alcuna procedura di tracciabilità e documentazione prevista dalle norme GMP o da altri standard aziendali.

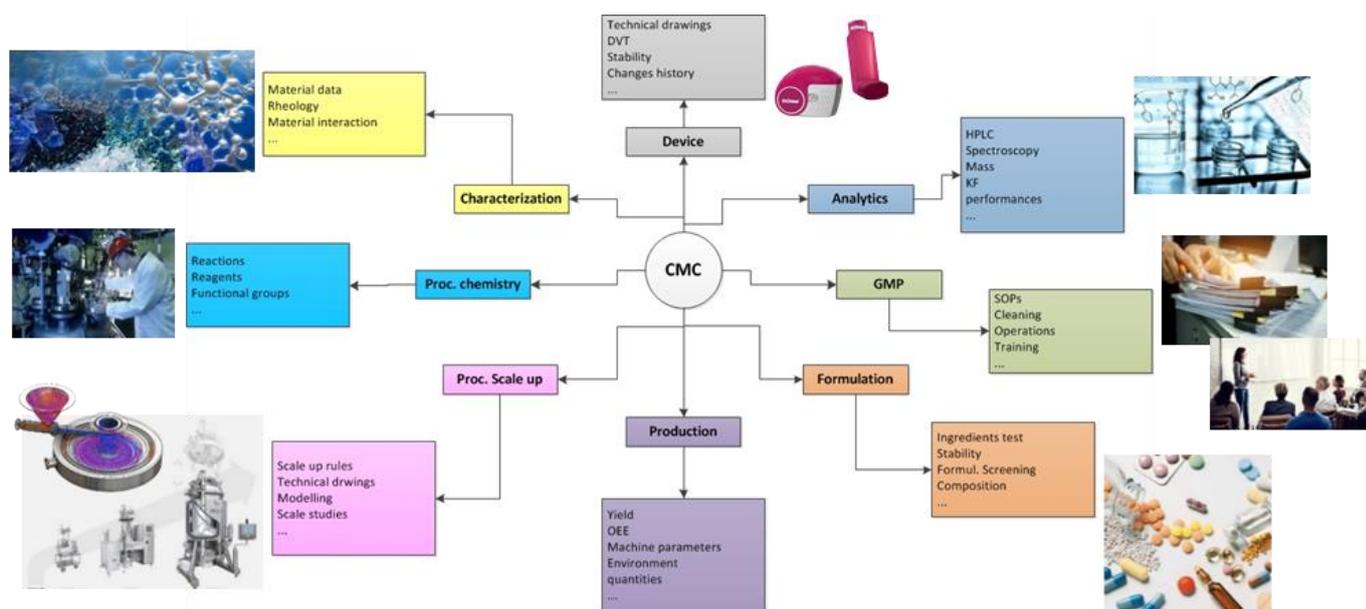


Figura 5: Eterogeneità dei dati CMC

I materiali non codificati sono sostanze il cui ciclo di vita non viene gestito dai sistemi gestionali come SAP o simili, siano essi materiali scaricati da magazzino o campioni acquistati come non codificati, manipolati e rielaborati in varie fasi e/o dipartimenti.

Le informazioni legate a ogni fase sono frammentarie e in alcuni casi ciascuna area utilizza un proprio sistema di informatica per caricare dati relativi allo stato del materiale.

Il problema nasce dal fatto che ogni operatore associa allo stesso materiale un dossier (per esempio un file) con nome differente non connesso con gli altri: non vi è quindi il collegamento con le informazioni generate durante le varie fasi del ciclo di vita del materiale. Emerge così la questione dei *silos* descritto nei capitoli precedenti.

Prendiamo ad esempio un materiale X, una volta scaricato da magazzino o altresì non codificato e quindi non gestito tramite SAP, può essere lavorato e analizzato; a seguito di queste operazioni i dati relativi alla sostanza vengono caricati nel sistema A con il codice 12345 insieme ad altre caratteristiche d'interesse di quella fase di processo. Il materiale procede attraverso le lavorazioni delle varie fasi fino a raggiungere un ipotetico STEP3 in cui subisce un'ulteriore trasformazione: a loro volta le relative analisi verranno inserite in un sistema informativo diverso da quello precedente (ad esempio B) con codice 13462 e metadati diversi rispetto a quelli registrati nelle fasi precedenti (Figura 6).

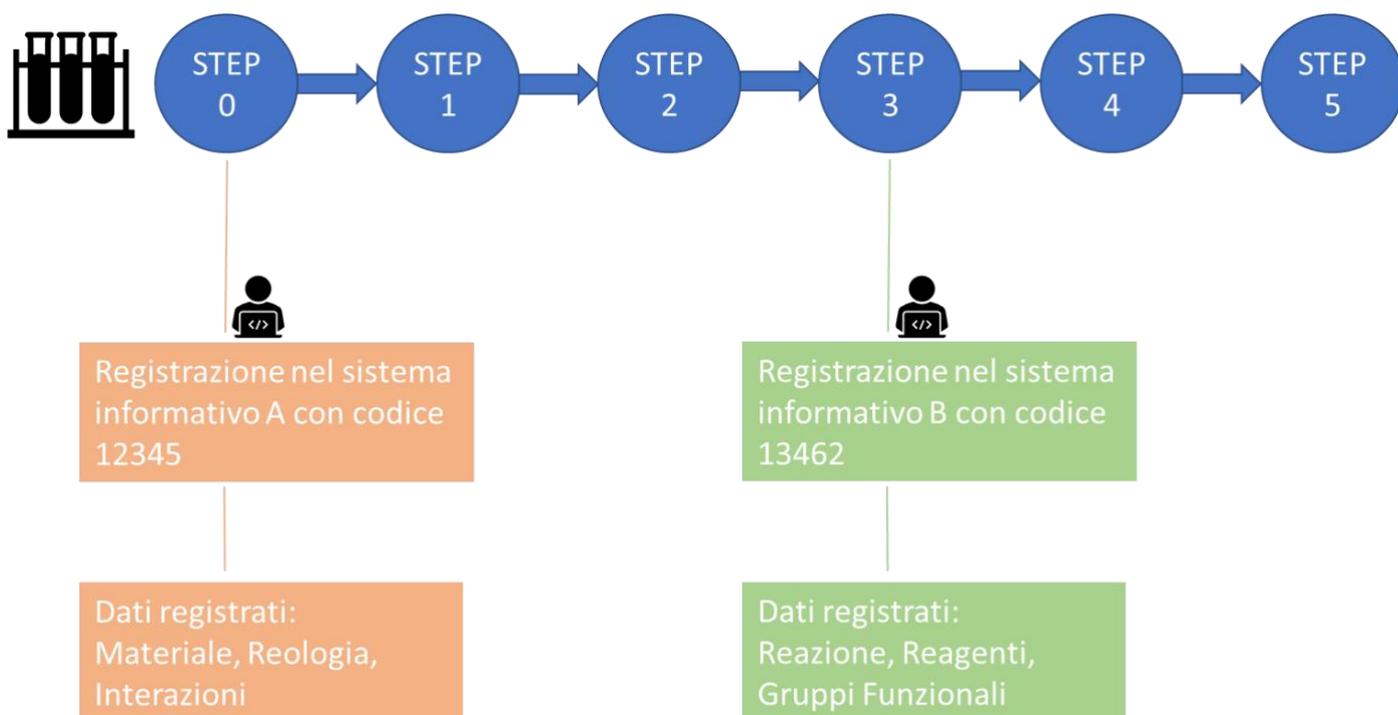


Figura 6: Esempio di problema di tracciabilità

Ogni linea di business raccoglie i dati in base agli utilizzi che ne deve fare e agli obiettivi che deve perseguire. Non essendo presente una chiave in grado di legare i codici generati nei vari step, seppur riferiti allo stesso materiale, risulta difficile l'identificazione delle connessioni facendo così emergere il problema dei silos.

La richiesta è stata dunque quella di riuscire a legare le informazioni, di uno stesso prodotto, generate in fasi diverse del ciclo di vita situate nei vari silos.

Per quanto riguarda tutto ciò che deve rispondere alle GMP, questo viene normalmente registrato e tracciato sui sistemi di gestione informatica aziendali validati per uso farmaceutico (ovvero "compliant" con le norme GxP).

Dipartimento CMC

Per produrre in modo appropriato un prodotto farmaceutico, è necessario definire processi di produzione specifici, caratteristiche del prodotto e *product testing* al fine di garantire che il prodotto sia sicuro, efficace e coerente tra i lotti. Queste attività sono note come *chemistry, manufacturing and control*.

Tutte le fasi evolutive del ciclo di vita del farmaco, dopo la scoperta del farmaco, coinvolgono la CMC. Durante lo sviluppo preclinico del farmaco, selezionato il *lead candidate* viene messa a punto la *route of synthesis* su scala di pochi grammi e vengono settati i metodi analitici appropriati per monitorare il prodotto.

Contemporaneamente vengono avviati i test preliminari di stabilità, si determinano le proprietà fisico-chimiche del prodotto, vengono scelte e testate le materie prime. Quando il processo di

sviluppo del farmaco entra nella fase clinica si procede con l'ottimizzazione e *scale-up* di processo, con una più accurata caratterizzazione del prodotto farmaceutico e con studi di stabilità ICH¹¹. In questa fase è richiesta la validazione dei metodi analitici con richieste più stringenti man mano che si procede con le fasi cliniche. La produzione dei lotti per i test clinici deve essere prodotta in accordo alle CGMP (*Current Good Manufacturing Practices*). La convalida di processo, di norma eseguita in concomitanza alla "fase clinica III" o comunque prima del *market approval*, garantisce la robustezza e riproducibilità di processo e quindi, in ultima analisi, che il prodotto soddisfi le stesse specifiche settate, ovvero, le stesse performance del farmaco testato negli studi clinici. Dopo che il processo di produzione è stato qualificato il farmaco può essere venduto, ma per il concetto di *continuous improvement* il prodotto entra nella catena della *life-cycle management* con un monitoraggio continuo per tutto il ciclo di vita del farmaco.

La CMC è una parte fondamentale della sperimentazione clinica del farmaco: l'approvazione del farmaco da parte delle autorità regolatorie del paese in cui vengono effettuati gli studi clinici potrebbe essere negata nei casi in cui la qualità del prodotto e del processo di fabbricazione non risultino in accordo alle GMP. Inoltre, le autorità di regolamentazione devono avere la certezza che vi sia coerenza tra il prodotto testato durante i test clinici e i lotti di prodotto commercializzati anni dopo.

L'ICH e la FDA hanno fornito una serie di guide per aiutare le aziende farmaceutiche con le aspettative CMC dei loro prodotti. In aggiunta, per produrre sia i lotti per le sperimentazioni cliniche che per il commercio, le aziende farmaceutiche italiane devono ottenere l'approvazione alla produzione da parte di AIFA¹².

Soluzione

È stata chiesta dunque una soluzione che fosse in grado di unificare e integrare le diverse fonti di dati e che fosse attuabile ottimizzando i tempi e i costi sia nella fase di ricerca sia nel lungo periodo con l'effettiva applicazione della nuova metodologia.

Il nuovo metodo, in aggiunta, avrebbe dovuto seguire il trend dell'azienda volto alla digitalizzazione dei processi in ottica dell'industria 4.0¹³

Per prima cosa è stato fatto un esercizio di mappatura delle fonti di dati e delle chiavi di connessione tra le informazioni che coinvolgono il reparto CMC.

Per fare ciò sono stati creati dei data model (usati per l'organizzazione, la gestione e l'archiviazione), per tutti i dati generati dalla CMC, con il fine di trovare delle chiavi di connessione tra di essi. Le chiavi hanno un ruolo fondamentale nel processo di estrazione dei dati, forniscono infatti il ponte che collega dati eterogenei raccolti su sistemi diversi. Più connessioni tra diversi sistemi sono disponibili, più dati possono essere interrogati per estrarre informazioni più approfondite.

¹¹ International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (www.ich.org)

¹² Agenzia Italiana del Farmaco (www.aifa.gov.it)

¹³ Il termine Industria 4.0 indica una tendenza dell'automazione industriale che integra alcune nuove tecnologie produttive per migliorare le condizioni di lavoro, creare nuovi modelli di business e aumentare la produttività e la qualità produttiva degli impianti.

Dall'analisi sono emersi i seguenti cinque gruppi, all'interno dei quali sono state individuate le fonti principali di dati con il quale il dipartimento si interfaccia:

- Streaming di dati da sistemi di automazione (per es. OSI-PI, DESIGO)
- Sistemi legacy (per es. LIMS, SAP, Adiuto, Hermes, Dikesi, Trackwise, Planisware)
- Cartelle condivise - sia archiviazione interna che archiviazione cloud - (per es. Sharefile, cartelle Box, unità di archiviazione di rete aziendale)
- Attrezzatura o file di strumenti
- Documenti cartacei

Nell'ambito dell'esercizio di mappatura, è stato chiesto di immaginare un'ipotetica situazione futura in cui ogni processo di business è stato completamente digitalizzato, rendendo automatiche tutte le connessioni dati e garantendo l'accesso alle informazioni correlate, che ora appaiono completamente scollegate.

Come emerso durante un'analisi preliminare svolta dall'azienda nel 2018 e confermato dall'esercizio di mappatura dei dati, molte fonti non possono essere collegate utilizzando le chiavi esistenti.

BI Code

L'idea è stata quella di assegnare ad ogni entità un metadato che fungesse da chiave universale (tecnicamente legato al concetto *foreign key*), che permettesse dunque di identificarla in modo univoco. Si è deciso quindi di creare un codice alfanumerico, denominato "**BI code**", una "carta d'identità" per l'informazione da archiviare, che permetta ex-ante di poter raggruppare dati tra loro collegati.

Questo codice fungerà da link trasversale per unire i dati di una stessa entità all'interno di *silos* (in questo caso i sistemi informativi) differenti (Figura 7).

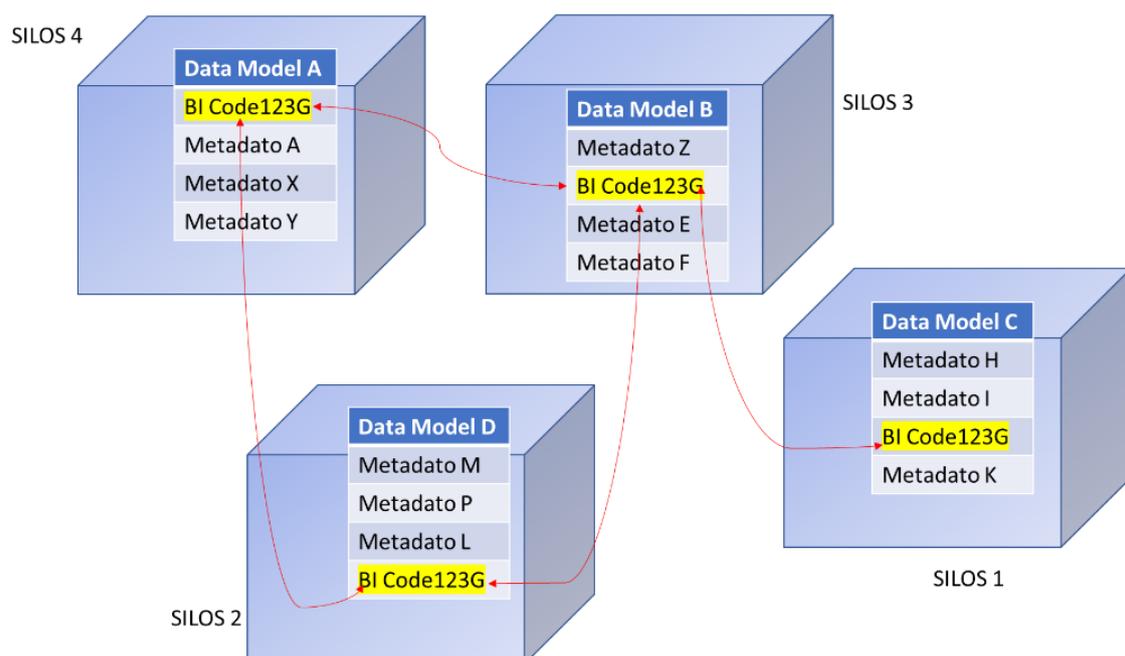


Figura 7: Esempio di collegamento attraverso BI Code

Metodologia di Approccio al Progetto

Bottom-up

Questo metodo è reso efficiente dall'utilizzo di tecniche e strumenti tecnologici scalabili in caso di variazioni notevoli della mole o della tipologia dei dati trattati.

L'approccio bottom-up va contro le metodologie tradizionali di gestione dei progetti dove il top management stabilisce obiettivi, fornisce linee guida, informazioni e piani attraverso processi estremamente formali. Nell'approccio top-down si sono evidenziate debolezze, poiché, accentrando il potere, le decisioni progettuali vengono influenzate dalle emozioni e dalle opinioni di poche persone, mentre il resto del team avverte di non avere voce in capitolo. Questo genera scarsa collaborazione, demotivazione e alienazione dal lavoro di gruppo.

L'accentramento della gestione può inoltre causare rigidità e rallentamenti sulla chiusura del progetto.

Al contrario, l'approccio bottom-up richiede l'intervento del team nel processo esecutivo di progetto. Chi lavora è invitato a partecipare a ogni fase del processo stimolando i membri del team a pensare creativamente e a individuare soluzioni.

Pur conoscendo i limiti di questo approccio, specialmente circa mancanza di chiarezza e controllo del progetto, si è scelto di adottarlo dopo aver analizzato:

- a) dimensioni del progetto - si è deciso di partire da uno use case che verrà illustrato nei paragrafi successivi per risolvere i problemi di un'area aziendale, per poi sviluppare una soluzione che in seguito a modifiche potrà essere utilizzata in più aree funzionali del centro ricerche. Eventuali modifiche non richiedono sforzi rilevanti grazie all'utilizzo di strumenti tecnologici scalabili in caso di rilevanti variazioni nella mole o nella tipologia dei dati trattati.
- b) il budget - si è partiti da un progetto privo di budget, l'obiettivo è quello di ottenere finanziamenti da parte dell'azienda dopo aver testato un prototipo di soluzione e averne evidenziati i vantaggi

POC e Prototipo

Per implementare la soluzione allo scenario proposto si è partiti da una breve fase iniziale di *proof of concept (POC)* nel quale è stata analizzata la fattibilità del progetto, con le risorse e le tempistiche a disposizione. Una volta dimostrata la realizzabilità della soluzione si è passati alla fase di prototipazione. POC e prototipo sono termini spesso usati in modo intercambiabile, tuttavia hanno due significati diversi.

Un *proof-of-concept* è un piccolo progetto creato per verificare se una certa idea o teoria sul prodotto può essere implementata; una piccola ricerca che dà il via libera per andare oltre con lo sviluppo di un prodotto.

Un prototipo è una versione iniziale, un modello limitato, costruito o creato per convalidare un concetto o un processo, lo scopo è quello di provare la POC e di fornire le specifiche successive per un sistema reale e funzionante piuttosto che teorico. Creando un prototipo viene testato il design, l'usabilità e la funzionalità del prodotto.

Metodo Scrum e Agile Project Management

Durante l'intero lavoro sono state seguite le linee guida della metodologia *Scrum* della famiglia dell'*Agile Project Management*.

“L'Agile Project Management è il lavoro di energizzare, arricchire e mettere gruppi di progetto in condizione di rilasciare valore per il business in modo rapido e affidabile, coinvolgendo i clienti e adattandosi in maniera continua ai loro bisogni e ai contesti in evoluzione¹⁴”.

In un contesto sempre più dinamico come quello attuale, è indispensabile la capacità di evolversi ed adattarsi velocemente ai cambiamenti. Nell'ultimo decennio, le tecniche tradizionali di Project Management si sono rivelate spesso inefficaci nel garantire il successo sia in termini di tempistiche che di costi e qualità, di una grande quantità di progetti indipendentemente dalla loro effettiva complessità. A causa di ciò vi è stata una spinta verso l'utilizzo e la diffusione di nuove metodologie per lo sviluppo, l'amministrazione e la gestione dei progetti e dei team impegnati nella loro realizzazione. Queste metodologie hanno come obiettivo quello di incrementare la produttività e l'efficienza del gruppo di lavoro, mediante l'introduzione di un forte cambio di paradigma nella gestione delle risorse di progetto, nella struttura organizzativa, nei piani strategici e, soprattutto, nella cultura aziendale.

L'Agile Project Management si inserisce nello scenario fortemente dinamico dell'attuale contesto di mercato, garantendo velocità e prontezza di risposta al cambiamento.

La figura 8 sottostante illustra brevemente alcuni dei principi fondamentali di Agile che sono stati formalizzati e racchiusi all'interno del Manifesto Agile¹⁵: l'attenzione al consumatore e al soddisfare i requisiti, l'approccio flessibile e positivo ai cambiamenti, le iterazioni e le frequenti consegne al cliente, la collaborazione quotidiana tra i membri del team, l'attenzione al lavoro di squadra e alla comunicazione, l'auto-organizzazione del team, il miglioramento continuo e la misura del progresso basata sulle caratteristiche del prodotto. Durante il progetto ci si è concentrati soprattutto su alcuni di questi principi, senza però trascurare gli altri.

¹⁴Augustine S. Managing agile projects: Prentice Hall PTR; 2005.

¹⁵ Manifesto for Agile Software Development: Kent Beck, Robert C. Martin, Martin Fowler; 2001

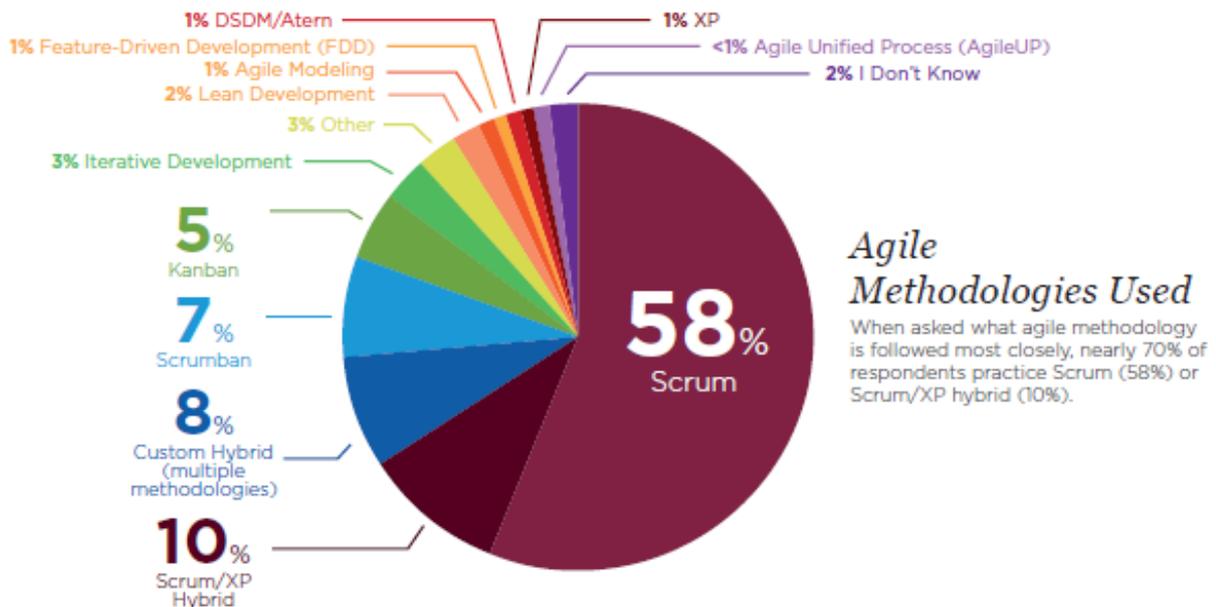
12 Principles of Agile



Figura 8: 12 Principi di Agile Project management

Per implementare i concetti espressi dall'Agile Project Management sono state introdotte numerose metodologie. Per il lavoro svolto si è deciso di seguire le linee principali della tecnica *Scrum* che, come mostrano i risultati del questionario dell'“Annual State of Agile Report” di VersionOne, risulta ampiamente essere la più adottata sul mercato, grazie soprattutto alla sua flessibilità e semplicità d'impiego.

AGILE METHODS AND PRACTICES



Questa tecnica viene introdotta per la prima volta nel 1986 da Hirotaka Takeuchi¹⁷ e Ikujiro Nonaka¹⁸ che la descrivono come una “strategia flessibile e olistica allo sviluppo di un prodotto, dove il team di sviluppo lavora come un’unica entità per raggiungere un obiettivo comune, in opposizione all’approccio sequenziale delle metodologie tradizionali”, questo approccio avrebbe accresciuto la flessibilità e la velocità nello sviluppo dei nuovi prodotti in commercio.

Questa tecnica, permette di andare a ridurre il rischio di insuccesso nello sviluppo software mediante la suddivisione temporale del progetto in iterazioni chiamate *sprint*, ovvero brevi cicli di sviluppo della durata di poche settimane. Ogni iterazione viene trattata come un piccolo progetto autonomo che deve contenere al suo interno un piccolo avanzamento nelle funzionalità del software. Durante gli *sprint* vengono costantemente implementate, valutate e aggiustate le diverse variabili di progetto, al fine di raggiungere la flessibilità necessaria a rispondere ai cambiamenti imprevisti.

Il team agile ingloba non solo chi si occupa di programmazione, ma anche tutti i portatori di conoscenza e tutti i soggetti in grado di definire come il prodotto dovrà essere fatto; prima dell’inizio di ogni *sprint*, è stata svolta una riunione di pianificazione in cui sono stati identificati gli obiettivi e i punti prioritari da completare durante l’iterazione. È stata dunque fondamentale la continua interazione con gli *stakeholder* per avere un quadro completo sulle esigenze da soddisfare e sulle modifiche da apportare nelle varie fasi dello sviluppo. Le specifiche del progetto vengono interpretate come elementi dinamici, che cambiano in linea con le necessità degli utenti durante l’avanzamento del progetto.

Al termine di ogni *sprint* si è tenuto un *Review Meeting*, a cui hanno partecipato sia il Team di Sviluppo che gli *stakeholders* con l’obiettivo di ispezionare l’incremento e di definire le successive attività da svolgere.

Lo Sprint Review Meeting è un incontro informale e la presentazione dell’incremento ha lo scopo di suscitare commenti e promuovere la collaborazione tra il Team di Sviluppo e gli *stakeholders*; la comunicazione diretta viene privilegiata rispetto a quella formale e cartacea tipica della gestione tradizionale dei progetti.

¹⁶ Annual State of Agile Report, VersionOne; 2010

¹⁷ https://en.wikipedia.org/wiki/Hirotaka_Takeuchi

¹⁸ https://en.wikipedia.org/wiki/Ikujiro_Nonaka

Use Case

Il caso d'uso a cui ci si è ispirati per la creazione del prototipo riguardava i problemi di tracciabilità dei prelievi di materiale che vengono inviati in CMC per effettuare analisi e test e dei successivi campionamenti effettuati nel laboratorio.

Nel momento in cui l'operatore di laboratorio fa richiesta, tramite apposito sito web, di un campione ad esempio di 100 grammi di un materiale X dal magazzino, il dipendente che si occupa del *sampling* effettua il prelievo senza annotare nel contenitore nessuna particolare informazione se non la quantità, il tipo di materiale e il lotto.

Tutti i lotti di magazzino sono registrati su ERP¹⁹, tuttavia colui che effettua il prelievo non indica sul campione nessun codice associato al sistema.

Se il chimico non vuole perdere informazioni riguardo alla data del prelievo, codice SAP e molti altri dati omessi, deve segnarsene su fogli cartacei o file personali.

Questi 100 g una volta giunti in laboratorio subiscono diverse lavorazioni e sotto-campionamenti, tutte queste devono sempre essere annotate autonomamente da parte dell'operatore cercando di far riferimento al campione principale: ad esempio 50 g vengono spediti ad un fornitore esterno, mesi dopo altri 20 g vengono utilizzati come reagenti di una reazione registrata su un ulteriore software, è importante che l'operatore riesca a ricondurre questo sub-sample al campione di 100g prelevato diverso tempo prima, operazione non semplice se i dati sono stati smarriti o se il secondo prelievo è stato fatto da un operatore non a conoscenza delle informazioni a monte.

Ulteriore problema sorge nel caso in cui un chimico si trova di fronte a due prelievi di uno stesso materiale di cui non è stata annotata la data di campionamento. L'operatore non solo non conoscerà da quanto tempo il materiale è in laboratorio, ma neanche come è stato conservato, ad esempio se a temperatura ambiente sul banco oppure in frigorifero.

I materiali che transitano in laboratorio possono essere addirittura campioni gratuiti, rappresentativi del lotto, richiesti al supplier per fare degli use test. Le informazioni relative a questi sample non passano attraverso nessun software aziendale, nel caso queste non fossero segnate dall'operatore di laboratorio non sarà possibile reperirle se non attraverso lunghe ricerche tra le e-mail e quindi causando perdite di tempo per i lavoratori.

L'operatore, non potendo rendere pubbliche le informazioni, diventa una sorta di silos di dati; la richiesta è stata dunque quella di riuscire a velocizzare le operazioni di ricerca, creare un link tra i vari campioni e tenere traccia e rendere accessibili, a tutti coloro che potrebbero farne richiesta, dei dati relativi ai materiali ad oggi non archiviati.

¹⁹ Enterprise Resource Planning

Soluzione

La soluzione proposta consiste nell'introduzione di una nuova metodologia di gestione dei materiali basata sull'etichettatura e la generazione di un codice BI da associare all'etichetta.

Nonostante l'introduzione di un codice BI garantisca diversi vantaggi, la sua implementazione richiede un notevole impegno da parte di tutte le unità coinvolte nel processo di generazione dei dati. In particolare, l'introduzione del BI Code richiede l'implementazione di nuove interfacce che verranno utilizzate per registrare i dati e generare i relativi BI Code.

Per fare ciò è stata creata appositamente un'applicazione web, in grado di registrare i dati, creare e leggere le etichette dei materiali e tracciare l'intero ciclo di vita del materiale e dei componenti che lo costituiscono. Per tenere traccia delle informazioni generate dalle azioni, ogni volta che si ottiene un campione o un sotto-campione viene generato un codice BI da associargli: quest'ultimo permette di collegare il *sub-sample* alle analisi corrispondenti registrate sui diversi sistemi aziendali. Lo storico dei dati, raffigurato attraverso alberi genealogici, permetterà di avere dei link tra i BI Code dei *codici padre* con quelli dei *codici figlio*.

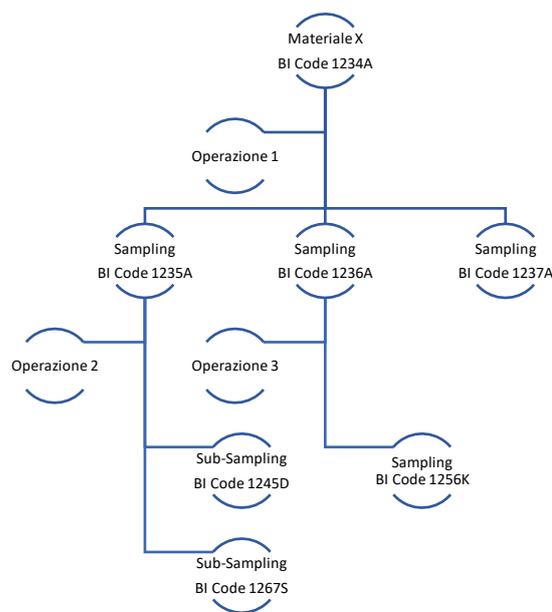


Figura 10: Esempio di BI Code chart

Le operazioni raffigurate in Fig.10 sono attività che possono sia modificare lo stato del materiale, cambiandone semplicemente la descrizione, sia operazioni che collegano più entità, come ad esempio una reazione chimica dove da diversi codici padre si ottiene un unico prodotto.

Web Application

Un'applicazione *Web* è un software applicativo eseguito su un Web server. Le applicazioni sono accessibili dall'utente tramite un browser Web con una connessione *Internet* attiva. Queste applicazioni sono programmate utilizzando una struttura *client-server*, ovvero un'architettura di

rete nella quale genericamente un computer client o terminale si connette a un *server* per la fruizione di un certo servizio. (Fig.11)

Il *client* è generalmente di bassa complessità, in quanto si limita ad operare come interfaccia verso il *server*. In informatica il termine client indica una componente che accede ai servizi o alle risorse di un'altra componente, detta *server*.

Dall'altro lato, il server, oltre alla gestione logica del sistema, deve implementare tutte le tecniche di gestione degli accessi, allocazione e rilascio delle risorse, condivisione e sicurezza dei dati o delle risorse.

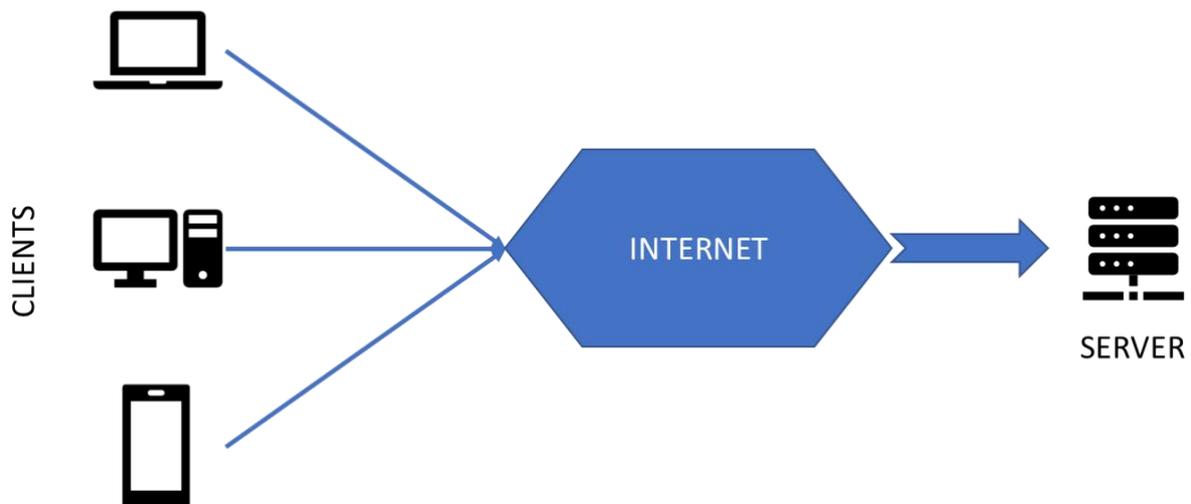


Figura 11

La creazione di una *Web application* è un'operazione alquanto complessa per l'azienda, tanto da spingerla a rivolgersi a fornitori esterni. Oltre alla lunga attesa per lo sviluppo sono richieste le competenze di un team di informatici e la continua interazione tra questi e gli operatori dei reparti che andranno a utilizzare l'applicazione in modo da customizzarla in base ai loro bisogni.

La scrittura del codice è solitamente svolta utilizzando linguaggi di programmazione adeguati, come ad esempio JavaScript per il front-end o PHP per il back-end.

Si è deciso di creare l'applicazione utilizzando una piattaforma low-code/no-code che - rispetto alle metodologie tradizionali di sviluppo software - offrono all'azienda di creare autonomamente la propria applicazione riducendo considerevolmente i costi rispetto alle metodologie tradizionali.

[Low-code/No-code Platform vantaggi e svantaggi](#)

Negli attuali ambienti IT, sta crescendo sempre di più la necessità di accelerare lo sviluppo del software a livello *enterprise* e i tool di sviluppo *low-code* rappresentano una delle nuove possibili soluzioni per soddisfarla. Quest'aspetto si riflette anche sulla competitività delle imprese dove le dinamiche di mercato si evolvono con rapidità e impongono ai responsabili del business d'innovare

prodotti e servizi aumentando l'agilità dell'organizzazione e comprimendo il più possibile il time-to-market.

Per capire il significato del concetto low-code, e quali caratteristiche distintive debba avere una piattaforma tecnologica improntata su questo paradigma, si può cominciare con la definizione fornita dalla società di ricerche Gartner: "una piattaforma in grado di supportare il rapido sviluppo e deployment delle applicazioni, attraverso l'utilizzo di tecniche low-code e no-code, come ad esempio la progettazione e lo sviluppo applicativo di tipo dichiarativo basato su modelli (*model-driven*), assieme all'implementazione semplificata ("*one-button*") delle applicazioni stesse".

In sintesi, grazie alle tecniche low-code, per lo sviluppo di un'applicazione non è più necessaria la scrittura di codice, basta una descrizione di come deve essere fatta l'applicazione e di quali operazioni deve svolgere. Tutto questo è reso possibile inserendo nell'interfaccia della piattaforma low code, attraverso operazioni "drag and drop", strumenti grafici, tecniche visuali e dichiarative. È l'intelligenza dell'algoritmo "dietro le quinte" che si preoccupa successivamente di tradurre il tutto in codice, rendendo così lo sviluppo più semplice e veloce.

La peculiarità di queste piattaforme di consentire all'utente di implementare in modo semplice i requisiti dell'applicazione, e di usare strumenti grafici e *building block* riutilizzabili per crearne di nuove, rende in una certa misura il processo di sviluppo SW accessibile anche ai non esperti di programmazione.

Dunque, le piattaforme *low-code* non soltanto aiutano gli sviluppatori professionisti ad accelerare la velocità di programmazione (John Rymer di Forrester ha scritto che il low-code offre il potenziale per "*rendere lo sviluppo del software fino a 10 volte più veloce rispetto ai metodi tradizionali*"), ma estendono l'adozione dei propri strumenti anche agli esperti del business che, senza necessità di scrivere codice, diventano in grado di dare il loro contributo e partecipare allo sviluppo dell'applicazione *enterprise*.

Nasce la figura del *citizen developer*, ovvero quell'utente di business (non esperto IT) in grado di sviluppare applicazioni utilizzando ambienti di sviluppo e runtime omologati dal dipartimento IT dell'azienda pur operando al di fuori di esso.

Questo nuovo profilo favorisce la strategia organizzativa aziendale del *job enrichment*, in quanto il lavoratore assume maggiori responsabilità, e del *job enlargement* arricchendo le attività svolte dal dipendente e aumentandone la motivazione.

Un'altra caratteristica del *low-code*, strettamente connessa alla velocità, è sicuramente l'agilità: la struttura "visuale" mette a disposizione un ambiente estremamente flessibile in cui l'interfaccia può essere personalizzata e adattata alle diverse necessità.

Inoltre, la possibilità di cambiare in tempi rapidi la struttura del programma permette un veloce adattamento dell'interfaccia a nuove esigenze o modifiche: si fornisce un'applicazione completa in poche settimane con possibilità di modificarla in pochi minuti.

Dal punto di vista dei costi un software personalizzato è difficile da mantenere per le aziende. La creazione dello strumento è solo l'inizio di un processo di supporto che dura per tutto il ciclo di vita del software. Spesso mantenere e manutene un prodotto o servizio è più laborioso e costoso che costruire la versione. Gli strumenti low-code – e più in generale tutti gli strumenti in cloud -

permettono di ammortizzare i costi di supporto in tutte le installazioni. Quando viene rilasciata una nuova versione del sistema operativo o una nuova norma sull'archiviazione dei dati, i fornitori la implementeranno una volta per tutti. È importante d'altra parte affidarsi a imprese, creatrici di piattaforme low-code e strumenti in cloud, affidabili e stabili.

In generale per quanto riguarda i costi di sviluppo, una variabile da tenere in considerazione è la complessità del sistema che dev'essere sviluppato, che spesso è determinata da quella del corrispondente processo di business. Una piattaforma low/no-code può permettere di sviluppare rapidamente un prototipo o un'applicazione: a seconda di chi la utilizza (per esempio programmatore esperto oppure *citizen developer*) il livello e tipo di risultato varia. In ogni caso si può considerare che in media l'impegno di una piattaforma di sviluppo low-code porti a una riduzione dei costi di sviluppo di circa il 50% rispetto a uno sviluppo tradizionale: questo risulta sia dalla riduzione dei tempi di sviluppo che da quello dei costi di manutenzione.

Infine, a supporto della metodologia agile, della facilità e della velocità d'implementazione, la maggior parte delle piattaforme low-code consentono all'utente di organizzare e gestire la propria applicazione in cloud garantendo sempre massima sicurezza e privacy. Se si cambia cloud o si utilizza una nuova banca dati non è necessario riprogrammare. Anche gli strumenti visivi *client* utilizzati sono in cloud, ciò garantisce che l'applicazione sia immediatamente trasponibile e utilizzabile. Il cloud offre inoltre la possibilità a più sviluppatori o team di collaborare e lavorare contemporaneamente sullo stesso modello.

Cloud

Il *cloud computing* può essere definito, in breve, come uno strumento che permette agli utenti di accedere e utilizzare da remoto, funzionalità hardware e software via Internet.

Un servizio cloud ha tre principali caratteristiche:

1. L'utente non ha bisogno di acquistare, installare o gestire autonomamente gli hardware o i software. Il fornitore del servizio di hosting su cloud acquista, gestisce e mantiene gli hardware e i software all'interno di un data center di cui è proprietario o locatario.
2. L'utente ha la possibilità di avviare il servizio, di usufruirne giorno per giorno e di aumentare o diminuire il livello di utilizzo, con un'interazione minima, o addirittura nulla, con il fornitore del servizio. L'interazione con il servizio cloud avviene tramite una connessione Internet, in modalità self-service.
3. Il costo si basa sulla quantità di servizio utilizzato. Generalmente la fatturazione si basa sul tipo di abbonamento o sul modello di licenza scelto.

È opportuno osservare che il termine cloud non è sinonimo di Web: per i dipendenti di un'azienda, il servizio cloud sostituisce il software che andrebbe altrimenti installato sulla propria postazione di lavoro o che verrebbe eseguito dal data center aziendale. Se si utilizza un sito, come ad esempio Amazon, per l'acquisto di materiale, si utilizza il Web. Se invece si usufruisce di Amazon Simple Storage Service per archiviare dati (anziché salvarli su un disco rigido locale o su un server dell'ufficio) ci si avvale di un servizio in cloud.

BettyBlocks



Figura 12: Forrester Wave²⁰

Betty Blocks è una piattaforma i cloud di tipo low-code (con prospettiva no code) per lo sviluppo di applicazioni aziendali (prima di questo genere a essere utilizzata in azienda) dotata di un'interfaccia drag-and-drop.

Betty Blocks consente agli utenti di creare flussi di lavoro delle applicazioni tramite un'interfaccia visiva. Gli sviluppatori possono anche condividere i processi applicativi con altri utenti della piattaforma, consentendo loro di implementare gli stessi processi nelle proprie applicazioni.

Le app sviluppate con Betty Blocks possono essere utilizzate su dispositivi mobili, tablet o desktop e tutte le applicazioni hanno un'interfaccia di back-office standardizzata. La piattaforma si distingue dalle altre senza codice per la facilità di utilizzo da parte di chiunque, pur vantando capacità abbastanza potenti da creare applicazioni sicure, Betty Blocks la prima e unica piattaforma no-code ad avere certificazione ISO27001²¹, complesse e di livello aziendale.

²⁰ The Forrester Wave™: Mobile Low-Code Platforms for Business Developers, Q3 2018

²¹ Standard di sicurezza informatica

Betty Blocks esegue automaticamente gli aggiornamenti della piattaforma, mentre tutte le applicazioni utente sono in esecuzione. Gli utenti accedono sempre alla versione più recente senza dover mai migrare i dati.

Inoltre, sebbene non sia necessario scrivere alcun codice per sviluppare applicazioni complesse, Betty Blocks supporta anche integrazioni JavaScript o HTML per una maggiore personalizzazione.

Nel caso di Betty Blocks, un ulteriore vantaggio consiste nella modello di licenza: il calcolo del canone annuale si basa sul numero di strutture “di base”, quali per esempio web services, tabelle del database, utilizzate per costruire una nuova applicazione. Il numero di utenti che usano i sistemi sviluppati non influisce sui costi, questo significa che non è necessario pagare di più se gli utenti aziendali aumentano, fattore importante nell’ottica di espandere l’applicazione ad ulteriori aree aziendali.

Stampante

Come detto in precedenza la nuova metodologia si basa sull’etichettatura; per l’acquisto della stampante e delle etichette si è fatto riferimento a un’importante fornitore mondiale già cliente dell’azienda.

I requisiti richiesti per la stampante sono stati i seguenti:

- Costi contenuti: essendo in fase di prototipazione è sufficiente un modello base;
- possibilità di connessione tramite Bluetooth con smartphone aziendale;
- Compatibilità con il software creato
- etichetta adesiva di qualità medio bassa: il test per la validazione del metodo non durerà più di qualche mese

Fasi del processo

Nei seguenti paragrafi verrà descritta la nuova metodologia proposta per la gestione delle fasi di campionamento e ricerca di informazioni relative ai materiali sotto analisi.

In fase di prototipazione i codici BI non saranno costituiti da codici alfanumerici ma da numeri progressivi

Fase 0

In questo step iniziale l'operatore di laboratorio invia la richiesta di prelievo materiale al magazzino (Figura 13)

In fase prototipale si è deciso di mantenere la metodologia tradizionale, con la richiesta del campione che avviene attraverso il sito web del magazzino: in futuro questa azione potrebbe essere implementata attraverso app.

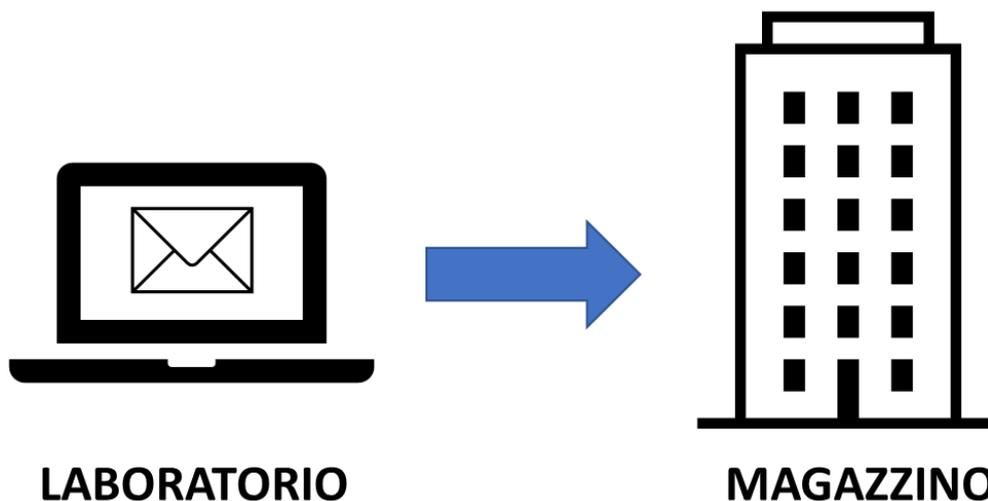


Figura 13: Fase 0

Fase 1

L'operatore di magazzino preleva il campione richiesto e apre l'applicazione attraverso smartphone aziendale.

L'utente potrà scegliere tra tre azioni:

1. Creare un nuovo BI Code da associare al campione
2. Aggiornare le informazioni relative a un materiale
3. Cercare informazioni relative a un materiale

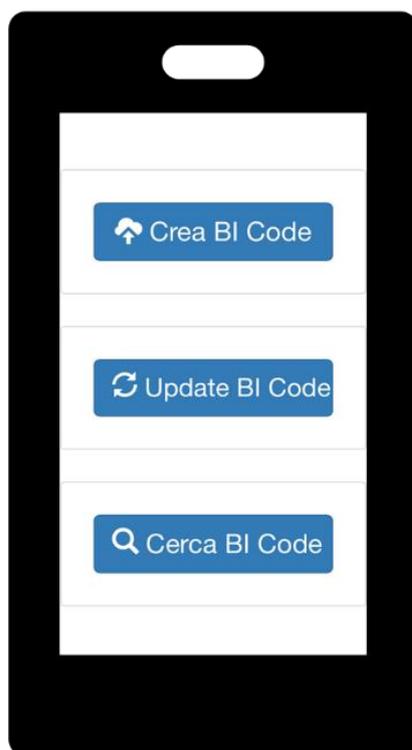


Figura 14: Scelta azioni possibili

Fase 2- Creazione BI Code

L'operazione di creazione di un nuovo BI Code può essere utilizzata da due attori diversi:

- l'operatore del magazzino nel caso in cui l'entità da codificare sia un campione prelevato da un lotto di materiale. Le sostanze presenti in magazzino sono tutte registrate nei sistemi informativi aziendali, in questo caso l'utente dovrà segnalare il numero del batch e il codice con il quale il materiale è presente nel sistema;
- l'operatore di laboratorio, nel caso in cui il materiale da codificare sia un sub-sample di un campione presente in laboratorio, quindi a cui era già stato associato un codice BI, oppure nel caso in cui il materiale sia il prodotto di più reagenti.

Dopo aver selezionato l'azione corrispondente, l'utente viene diretto a una pagina in cui va definito il software aziendale da cui è stato scaricato il materiale (in fase prototipale è stato scelto solo il sistema SAP) o se il campione prelevato è un sub-sample o il prodotto di reagenti a cui era stato associato un codice BI in precedenza.

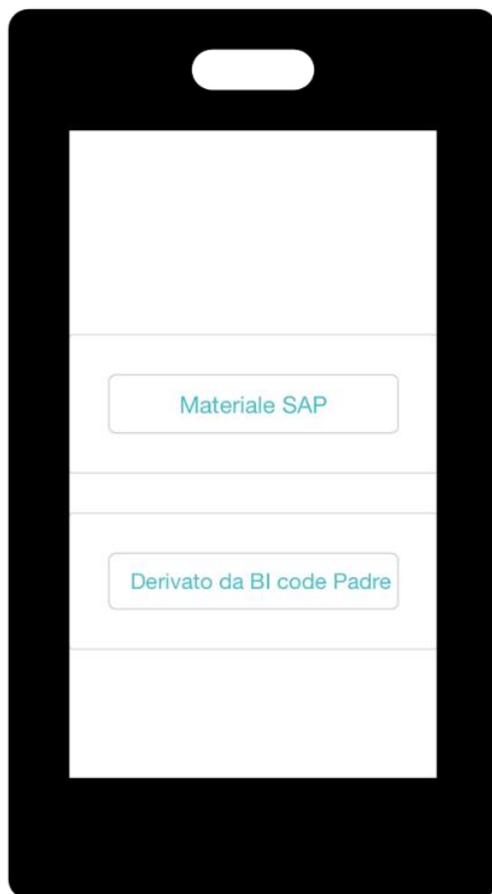


Figura 15: Fase 2

Fase 2.1- Materiale scaricato da software aziendale

In questa sezione l'operatore inserisce le informazioni relative al materiale prelevato (tra cui il tipo di materiale, il numero di lotto, la data del campionamento e il relativo codice collegato all'entità nel sistema informativo aziendale).

Automaticamente l'applicazione genererà la data di creazione della nuova istanza dell'oggetto che rappresenta il materiale.

Nonostante l'attività di mappatura dei metadati effettuata in precedenza, si è deciso, per il prototipo, di lasciare i campi d'inserimento sufficientemente liberi in modo che possano emergere eventuali metadati non rilevati.

SAP code

Batch

Materiale

Descrizione del contenuto

Crea

Figura 16: Fase 2.1

Fase 2.1.1

Una volta completata la compilazione del *form* precedente (Fig.16) premendo il tasto “Crea” l’applicazione genererà e assegnerà un codice BI al materiale.

L’utente viene indirizzato ad una pagina in cui sarà visibile il codice BI (la Fig.17 mostra il *BI Code* assegnato al materiale, in questo caso il “36”) e il QR code all’interno del quale sono racchiuse le informazioni relative al campionamento effettuato.

L’utente stampa su etichetta adesiva il QR Code e lo applica sul sacchetto contenente il campione

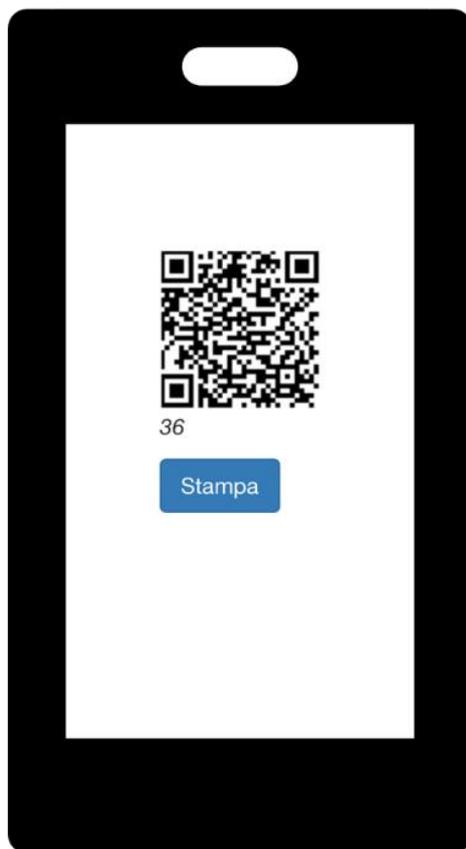


Figura 17: Fase 2.1.1

Fase 2.2- Materiale derivato da BI Code padre

Questa azione viene utilizzata dagli operatori di laboratorio nel caso in cui venga effettuato un sottocampionamento o sia da generare un codice BI per il prodotto di una reazione.

Dopo che l'utente ha inserito i metadati fondamentali per identificare l'elemento - sono stati scelti solo la tipologia di materiale e la descrizione (Fig.18) ma è possibile aggiungerne molti altri - l'applicazione genera il codice BI da associare e indirizza l'operatore alla pagina in cui devono essere inseriti i codici BI ("codici padre") dei reagenti (Fig.18: i codici padre sono "37, 38, 39"). Nel caso in cui il materiale in analisi sia un sub-sample ci sarà un solo codice padre associato in quanto non deriva dalla combinazione di più materiali.



Figura 18: Fase 2.2

Fase 2.2.1

Una volta terminata la compilazione dei campi l'utente potrà, come osservato precedentemente, stampare l'etichetta con il codice QR associato alle informazioni del materiale e utilizzarla.



Figura 19: Fase 2.2.1

Aggiornamento BI Code

Attraverso questa sezione, dopo aver cliccato sul tasto “Update” (Fig.14), l’operatore può modificare e aggiornare le informazioni relative al materiale che ha analizzato come ad esempio:

- i test effettuati
- la metodologia di conservazione
- la temperatura
- il codice delle analisi associate in un sistema aziendale

e molte altre informazioni che attualmente sono assenti o che vengono registrate in ambienti di archiviazione privati.

Sarà sufficiente inserire il codice BI del campione per essere reindirizzati nella pagina apposita dove l’operatore potrà inserire commenti sui cambiamenti ottenuti sul materiale in seguito ai test svolti.



Figura 20: Schermate Aggiornamento BI

Fase 3- Rintracciabilità

Qualora un utente abbia bisogno di ottenere informazioni relative al campione sarà sufficiente inquadrare con lo smartphone aziendale il codice QR per essere reindirizzato alla relativa info page (Fig.21).

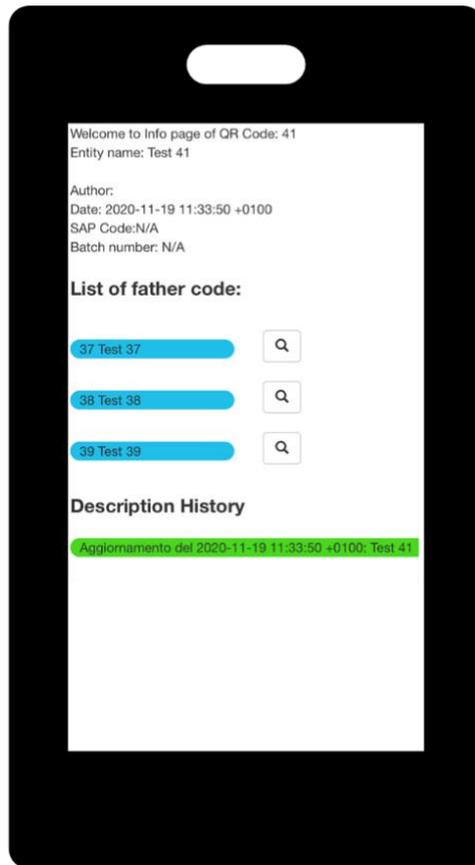


Figura 21: Esempio di Info page

Premendo sul simbolo della lente d'ingrandimento a lato di ogni codice padre si ottiene la pagina delle informazioni relative al determinato codice; è così possibile ripercorrere una sorta di albero genealogico che permette di tener traccia dell'intero ciclo di vita dei materiali.

In alternativa, sarà sufficiente che l'operatore entri nell'applicazione e attraverso l'opzione di ricerca (Fig. 14) definisca per quale metadato vuole filtrare l'analisi.

Il risultato non sarà, come nel caso precedente, un'entità definita, bensì una famiglia di item che rispondono ai criteri di ricerca impostati, sarà poi l'operatore a scegliere tra quelli elencati l'elemento di cui vuole ottenere informazioni.

L'aspetto relativo alla genealogia del materiale sarà importante dal punto di vista degli analisti in quanto, trovandosi di fronte a un prodotto difettoso, possono risalire allo step in cui si è verificato l'errore o al reagente che presentava caratteristiche non congrue a quelle standard.

Ad esempio, assumiamo che il prodotto con BI Code A, in seguito alle analisi effettuate e caricate sul sistema LIMS con codice 125, abbia riscontrato dei difetti.

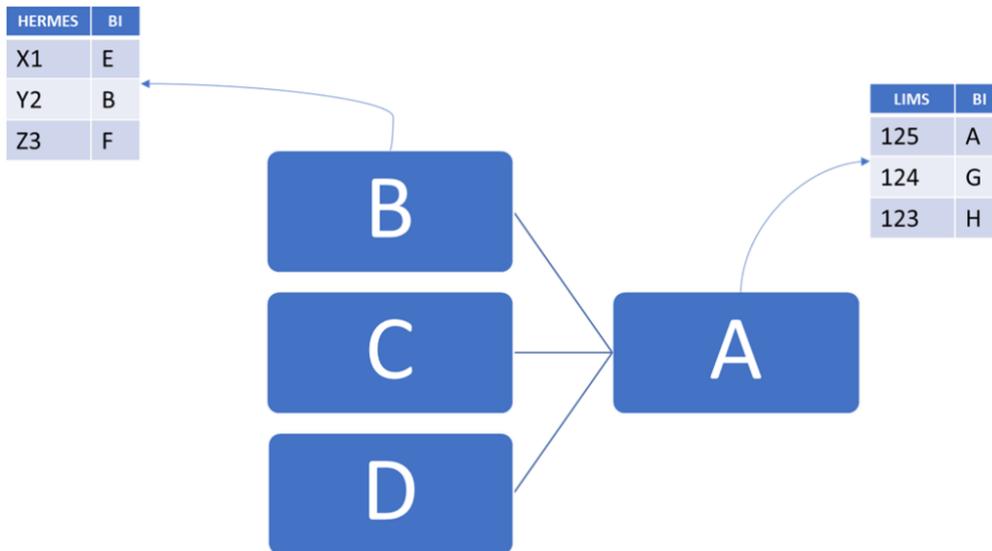


Figura 22: Genealogia del prodotto di una reazione

L'analista che vuole identificare l'origine del problema osservando la genealogia del prodotto potrà incrociare le analisi di A presenti su LIMS con quella dei suoi reagenti presenti su altri sistemi, nell'esempio l'operatore entrando dalla pagina informativa di A procede a ritroso fino a ottenere i codici delle analisi relative ai reagenti, in questo caso le informazioni relative a B sono situate su Hermes con codice Y2, e confrontando i dati o rivolgendosi al dipartimento che ha generato il dato potrà individuare la natura del problema. (Fig.23)

Welcome to Info page of QR Code: **A**
 Entity name:
 Author:
 Date: 2020-09-24 09:51:33 +0200
 LIMS Code: 125
 Batch number: N/A

List of father code:

- B [Search Icon]
- C [Search Icon]
- D [Search Icon]

Description History

- Aggiornamento del 2020-09-25 10:59:48 +0200: Temperatura X °C
- Aggiornamento del 2020-09-25 11:00:01 +0200: Prodotto non conforme

CLICCANDO QUI OTTENIAMO LA STESSA PAGINA MA RELATIVA AL CODICE BI «B». AL POSTO DEL CODICE LIMS SARÀ PRESENTE IL CORRISPETTIVO CODICE HERMES

Figura 23: Procedimento di ricerca informazioni

Analisi Economica

Sotto le ipotesi della creazione di un prototipo aziendale poco complesso e in ottica di un numero elevato di user si è deciso di stimare il costo dello sviluppo e manutenzione attraverso le tecnologie tradizionali doppio rispetto allo sviluppo per mezzo delle tecniche low-code.

Costo di sviluppo

Nel seguente scenario verranno stimati i costi di sviluppo confrontando due possibili figure a cui l'azienda potrebbe assegnare il lavoro:

1. Il dipendente aziendale citizen developer
2. Consulente esterno

Si ricorda che la piattaforma Betty Blocks, ha un calcolo del canone annuale basato sul numero di strumenti grafici utilizzati per la creazione dell'applicazione, al contrario di altre piattaforme che fanno pagare una tariffa fissa in base al numero di utenti; questo permette teoricamente di avere un numero infinito di utilizzatori di una stessa applicazione senza che il costo per l'azienda aumenti.

Assunzioni:

- Le applicazioni sviluppate dalle due figure utilizzano uno stesso numero di strutture (80) al costo di 75 €*anno/struttura.
- Costo sviluppatore esterno: 500 €/gg
- Costo citizen developer: 230 €/gg valore stimato moltiplicando lo stipendio medio nel settore farmaceutico per un fattore 2 che rappresenta i costi aziendali.
- Coefficiente velocità di sviluppo $\frac{\text{sviluppatore esterno}}{\text{citizen developer}} = \alpha$

Per il seguente progetto si è assunto un valore di $\alpha = 1.85$, valore che deve tener conto della complessità dell'applicazione che si vuole sviluppare e dall'esperienza del dipendente con lo strumento tecnologico.

- Giorni di lavoro sviluppatore esterno: 32 gg
- Giorni di lavoro citizen developer = $32gg * 1,85 = 60 gg$
- Non sono considerati nel calcolo i costi relativi al training per l'utilizzo dell'applicazione, si ipotizza siano stati ammortizzati in precedenza

Calcolo dei costi:

Poiché le due applicazioni da ipotesi presentano uno stesso canone annuale, vengono analizzati solo i costi da allocare alla figura che ha sviluppato l'applicazione.

$$\text{Costo sviluppatore esterno} = 500 \frac{\text{€}}{\text{gg}} * 32 \text{gg} = 16'000\text{€}$$

$$\text{Costo c.d.} = 230 \frac{\text{€}}{\text{gg}} * 60 \text{gg} = 13'800\text{€}$$

Conclusione: Sotto le ipotesi sopracitate, lo sviluppo di applicazione attraverso la figura del citizen developer ha un costo inferiore rispetto all'assegnazione del lavoro ad un consulente esterno se è rispettata la condizione del coefficiente $\alpha \leq 2,05$

Riesame Cost of not finding information

Nelle successive analisi si è andato a rianalizzare il primo scenario presentato dall'azienda IDC ipotizzando l'introduzione di un'applicazione che migliora la tracciabilità, alla ricerca di quel valore percentuale X di ricerche fallite al di sotto del quale l'azienda ottiene benefici dal punto di vista di una riduzione dei costi nello stesso anno in cui viene adottata la nuova tecnologia. A dimostrazione dei vantaggi della scalabilità dell'applicazione lo scenario verrà riproposto facendo il calcolo sull'utilizzo di 100, 500 e 1000 lavoratori.

Nell'analisi è stato effettuato il cambio di valuta euro-dollaro con tasso di cambio 1 a 1

Si ipotizza un costo annuale dell'applicazione di 15'000€ e si conservano le restanti ipotesi dello scenario presentato da IDC

Calcolo dei costi scenario:

$$\frac{80000 \frac{\text{€}}{\text{anno}} * \left(2,5 \frac{\text{h}}{\text{gg}} * 5 \frac{\text{gg}}{\text{week}} * 52 \frac{\text{week}}{\text{anno}} * X\% \right)}{2080 \frac{\text{h}}{\text{anno}}} * t \text{ worker} + 15'000 = z \frac{\text{€}}{\text{anno}}$$

$$t = 100, z = 1'250'000; X\% = 49,4\%$$

$$t = 500, z = 6'250'000; X\% = 49,88\%$$

$$t = 1000, z = 12'500'000; X\% = 49,94\%$$

Conclusioni: per ottenere benefici economici nel primo anno l'applicazione deve ridurre il numero di ricerche fallite di una percentuale di poco inferiore rispetto allo scenario classico. Si osserva inoltre come la differenza di percentuale con lo scenario tradizionale si riduce all'aumentare degli user, a favore della scalabilità dell'applicazione.

La riduzione di ricerche fallite significa ridurre il tempo di ogni singola ricerca e maggior tempo per effettuare ulteriori.

Nel caso t=100 sarà sufficiente avere un risparmio di circa 1 minuto e 48 secondi sulle analisi fallite per avere un beneficio sull'utilizzo dell'applicazione, nel secondo caso di 22 secondi e nel terzo di 11 secondi

Considerazioni

Commentando nello specifico lo use case preso in esame la nuova metodologia introdotta porta sicuramente al risparmio in termini di tempo e denaro evitando che vi sia una mancanza d'informazioni riguardo al campione, ciò porterebbe al dover ripetere il campionamento, le relative analisi e le manipolazioni coinvolgendo più soggetti all'interno dell'azienda e nel peggiore dei casi il blocco della produzione su un determinato lotto.

Tuttavia, il vantaggio che si ottiene con l'utilizzo della soluzione proposta non riguarda solo il tempo e il denaro risparmiato, quest'aspetto per un'azienda che si occupa di ricerca è maggiormente legato alla produzione industriale; il vero beneficio, all'interno di un dipartimento di ricerca e innovazione, risiede nel valore del dato, nel conservare ed avere determinate informazioni che portano ad

accrescere ciò che per un'azienda farmaceutica è fondamentale, ovvero la knowledge e il know-how.

Non perdere la tracciatura, l'origine, la storia di tutto ciò che riguarda un materiale va a beneficio della conoscenza. Senza di questo nuovo strumento alcuni dati e risultati possono essere facilmente smarriti.

Conclusioni

Questa nuova tecnica non solo permette di tenere traccia degli spostamenti effettuati dal materiale e della relativa genealogia, ma permette inoltre di ottenere e conservare informazioni che attualmente non vengono fornite a monte (dal magazzino) o che essendo annotate dagli operatori su fogli cartacei o documenti privati richiedono svariato tempo per la ricerca e nella peggiore delle ipotesi possono essere smarrite.

La digitalizzazione dei dati fa inoltre sì che le informazioni siano potenzialmente disponibili a tutta l'azienda, evitando la creazione di silos di dati (rappresentati dai computer dei singoli utenti), così che tutti i dipendenti del reparto abbiano un'ampia e soprattutto rapida accessibilità alle informazioni.

Oltre ai benefici citati in precedenza, l'utilizzo dell'applicazione potrebbe portare a far emergere nuovi vantaggi non considerati durante il progetto.

È importante osservare come l'applicazione non fornisca tutte le informazioni relative a un'entità ma solo quelle non registrate negli altri sistemi. Il BI code non contiene informazioni presenti in altri sistemi aziendali, ma permette di collegare due fonti di informazioni che senza di esso non avrebbero alcuna relazione evidente.

L'aumento della tracciabilità fa parte di quel processo necessario all'azienda per garantire elevati standard di qualità, permettendo di ottimizzare il lavoro dei dipendenti e di ridurre notevolmente il tempo dedicato alla ricerca di informazioni; permette altresì di evitare la ripetizione di attività già svolte o di effettuare test su materiali che hanno riscontrato difetti in step precedenti.

Il progetto ha permesso di risolvere le problematiche presentate in maniera innovativa e con una minimizzazione dei costi, il risultato però che ha assunto maggior rilevanza non è stato tanto il risparmio economico e di tempo che si ottiene grazie all'applicazione ma è la capacità di conservare informazioni che porta alla crescita della knowledge aziendale.

Lo sviluppo dell'applicazione è stato approvato dal management aziendale poiché in linea con la cultura aziendale e le prospettive di digitalizzazione dei processi.

Validazione e prospettive future

Per la validazione del prodotto si è stabilito un periodo di prova di circa tre mesi. In questo intervallo di tempo verrà osservato l'utilizzo da parte degli operatori e un'analisi accurata dei dati registrati con il fine di apportare modifiche sull'applicazione nel caso di colli di bottiglia e di stabilire i campi definitivi per le entità da registrare. Potrebbero inoltre emergere utilizzi non considerati a inizio progetto.

In seguito, verranno somministrati dei questionari agli utenti, dopo un'analisi primaria si otterranno feedback costruttivi che potranno portare al miglioramento del software e di conseguenza dell'attività di tracciabilità.

Una volta ottenuto il prodotto finito, potrebbe essere utile effettuare un'analisi economica più dettagliata per rilevare quantitativamente i vantaggi derivanti dall'utilizzo dell'applicazione; poiché il software compie attività mai svolte prima dai dipendenti potrebbero emergere ed essere utilizzati KPI diversi da quelli definiti ex-ante.

Infine, verrà valutata la scalabilità dell'applicazione per l'utilizzo in più dipartimenti, eventuali modifiche sull'applicazione non richiederanno ingenti sforzi economici in quanto la piattaforma no-code permette di effettuare cambiamenti in modo agile.

Appendice

Good Manufacturing Practice

Le Good Manufacturing Practices (**GMP**) o Norme di Buona Fabbricazione (**NBF**), sono un insieme di regole, procedure e linee guida in base alle quali vengono prodotti i farmaci (in Europa si usa il termine medicinali), i cibi e le sostanze farmacologicamente attive (in inglese *Active Pharmaceutical Ingredients*, **API**).

L'industria farmaceutica della Comunità Europea garantisce un livello elevato di qualità nelle fasi di sviluppo, fabbricazione e controllo di medicinali. Un sistema di autorizzazione alla immissione sul mercato garantisce che tutti i medicinali siano valutati da un'autorità competente al fine di assicurarne la rispondenza alle norme vigenti in materia di sicurezza, qualità ed efficacia, mentre un sistema di autorizzazioni alla fabbricazione garantisce che tutti i prodotti autorizzati e venduti sul mercato europeo vengano fabbricati unicamente da fabbricanti autorizzati, la cui attività è soggetta a regolari ispezioni da parte delle competenti autorità. L'autorizzazione alla fabbricazione è obbligatoria per tutti i fabbricanti di prodotti farmaceutici della Comunità europea, a prescindere dal fatto che i loro prodotti vengano collocati sul mercato comunitario o sul mercato estero.

Le due direttive che contengono i principi e le linee guida sulle norme di buona fabbricazione (GMP) dei medicinali sono state adottate dalla Commissione nel 1991: la prima riguarda i medicinali per uso umano (direttiva 91/356/CEE), la seconda quelli per uso veterinario (direttiva 91/412/CEE). Le relative linee guida sono state pubblicate nella guida per la buona fabbricazione dei medicinali e costituiranno parametro valutativo delle domande di autorizzazione alla fabbricazione e per le ispezioni condotte presso i fabbricanti di prodotti medicinali.

Tutti gli Stati Membri e il settore industriale stesso concordano sul fatto che i requisiti previsti dalle GMP sulla fabbricazione di medicinali veterinari sono gli stessi di quelli applicabili alla fabbricazione di medicinali per uso umano. Alcune modifiche particolari alle linee guida delle GMP sono descritte nei due allegati relativi ai prodotti medicinali veterinari e ai medicinali ad azione immunologica.

La guida è articolata in capitoli, ognuno dei quali introdotto da un principio generale. Il capitolo 1 sulla gestione della qualità descrive il concetto fondamentale di garanzia di qualità applicato alla fabbricazione dei medicinali; ognuno dei capitoli successivi, dopo il paragrafo del principio generale che delinea gli obiettivi di quel capitolo in tema di garanzia di qualità, fornisce ai fabbricanti sufficienti dettagli sugli aspetti fondamentali da prendere in considerazione per tradurre in pratica il suddetto principio. Oltre agli aspetti generali delle norme per la buona fabbricazione delineati nei nove capitoli della guida, una serie di allegati fornisce dettagli su specifici settori di attività.

La prima edizione della guida è stata pubblicata nel 1989 e comprendeva un allegato sulla fabbricazione dei medicinali sterili.

La guida non tratta il problema della sicurezza del personale addetto alla fabbricazione. Tale aspetto può infatti rivestire particolare importanza nella fabbricazione di alcuni medicinali ad esempio quelli altamente attivi, biologici e radioattivi. Esso rientra nel campo di applicazione di altre normative comunitarie o nazionali. Nella guida si presume che i requisiti di qualità, sicurezza ed efficacia di un medicinale previsti dalla relativa autorizzazione alla immissione sul mercato siano sistematicamente impliciti nel processo di fabbricazione e nella fase di controllo e di distribuzione da parte del titolare dell'autorizzazione alla fabbricazione.

La guida non intende limitare in alcun modo lo sviluppo di nuove idee o nuove tecnologie purché queste siano convalidate e forniscano un livello di garanzia di qualità per lo meno equivalente a quello fissato nella guida.

Una componente estremamente importante delle NBF è la necessità di documentare ogni aspetto del processo, ogni attività ed ogni operazione. Se la documentazione non è corretta e in ordine, se non evidenzia il modo in cui il prodotto è stato fabbricato e verificato, permettendo la rintracciabilità e il ritiro dal mercato in caso di problemi futuri, allora il prodotto viene considerato contaminato.

La versione delle NBF elaborate dall'Organizzazione Mondiale della Sanità viene utilizzata dalle industrie farmaceutiche e dalle autorità di controllo del settore in oltre 100 paesi nel mondo, in particolare nei paesi in via di sviluppo. Nell'Unione Europea sono in vigore le relative norme europee, più restrittive rispetto a quelle dell'OMS, mentre negli Stati Uniti è in vigore la versione della Food and Drug Administration (FDA, l'ente che presiede alla regolamentazione dei prodotti alimentari e dei farmaci), anch'essa con prescrizioni aggiuntive rispetto a quelle elaborate dall'OMS. Meccanismi analoghi alle NBF sono in uso in altre nazioni (Australia, Canada, Giappone, Singapore ecc.), sempre con prescrizioni molto articolate e sofisticate.

Negli Stati Uniti, la FDA diffonde le linee di condotta relativa alle NBF tramite il meccanismo del Registro Federale e numerose linee guida dirette all'industria. Le NBF negli Stati Uniti sono una combinazione di leggi (principalmente il CFR 21, parti 210 e 211), delle migliori e attuali modalità produttive industriali e della filosofia dell'FDA espressa nelle proprie pubblicazioni, guide e linee guida, e altra documentazione interna come le Guide alle Procedure di Conformità e alcune parti del Manuale di Ispezione delle Operazioni Produttive. Quindi le NBF sono soggette a un processo di continua evoluzione in relazione ai miglioramenti produttivi di ogni azienda del settore.

L'ambito di applicazione delle NBF statunitensi è definito nel regolamento FDA (21 CFR210.1) dove è stabilito che le regole citate nel suddetto capitolo devono essere considerate "NBF minime e correnti"; le regole stesse devono essere considerate "dinamiche" e costantemente passibili di variazioni correlate con l'emergere di nuove tecnologie, nuove concezioni o influenze. Per questo motivo le NBF statunitensi vengono definite "correnti" (o cNBF/CGMP). Contrariamente alle NBF europee e ad altre NBF che sono normalmente aggiornate e ri-edite ogni 5-7 anni, il testo delle US NBF è rimasto sostanzialmente invariato per 30 anni, visto che l'FDA non sente alcun'esigenza pressante di elaborare revisioni poiché rimane valido il concetto che le norme di base devono essere "minime e correnti".

Contrariamente a molti altri paesi, le delibere e le pubblicazioni dell'FDA, incluse le pubblicazioni interne, sono soggette alla Legge sulla Libertà dell'Informazione negli Stati Uniti e ciò comporta che quasi tutti i documenti succitati possano essere scaricati via Internet. Una risorsa chiave per le NBF applicate alla industria farmaceutica è il Portale WEB dell'FDA per l'Industria Farmaceutica.

Sebbene ogni NBF nel mondo includa concetti simili, ci sono comunque delle differenze importanti. Nelle NBF Europee, e nelle NBF di quelle nazioni che fanno parte del Piano di Cooperazione nelle Ispezioni Farmaceutiche (PIC/S), la responsabilità di assicurare che i farmaci siano prodotti in accordo con metodi di produzione e controllo registrati presso le autorità e nel rispetto delle NBF, fa capo alla Persona Qualificata (PQ o QP — *Qualified Person*), definita anche Persona Responsabile che si assume la responsabilità personale e legale di certificare che ciascun lotto di prodotto è stato

prodotto in accordo con le autorizzazioni commerciali e con le NBF. Negli Stati Uniti le responsabilità della PQ sono conferite nell'ambito del gruppo di Assicurazione Qualità della società.

In Europa e nei paesi membri del PIC/S, le responsabilità personali e legali per non-conformità con le autorizzazioni commerciali e con le NBF sono conferite alla PQ che rischia multe e pene detentive per la mancata conformità. Negli Stati Uniti la responsabilità finale, e le eventuali multe o pene detentive, sono conferite al *Chief Executive Officer* (CEO) della società.

Big Data

Sebbene sia un termine oggi molto utilizzato, il concetto di **big data** ha origini incerte. Nonostante vi siano riferimenti già da metà degli anni Novanta, la figura sottostante mostra come il termine abbia avuto un forte utilizzo solo a partire dal 2011, probabilmente grazie a iniziative promozionali di aziende leader nel settore tecnologico, come ad esempio IBM, che hanno investito nella costruzione del mercato dell'analitica.

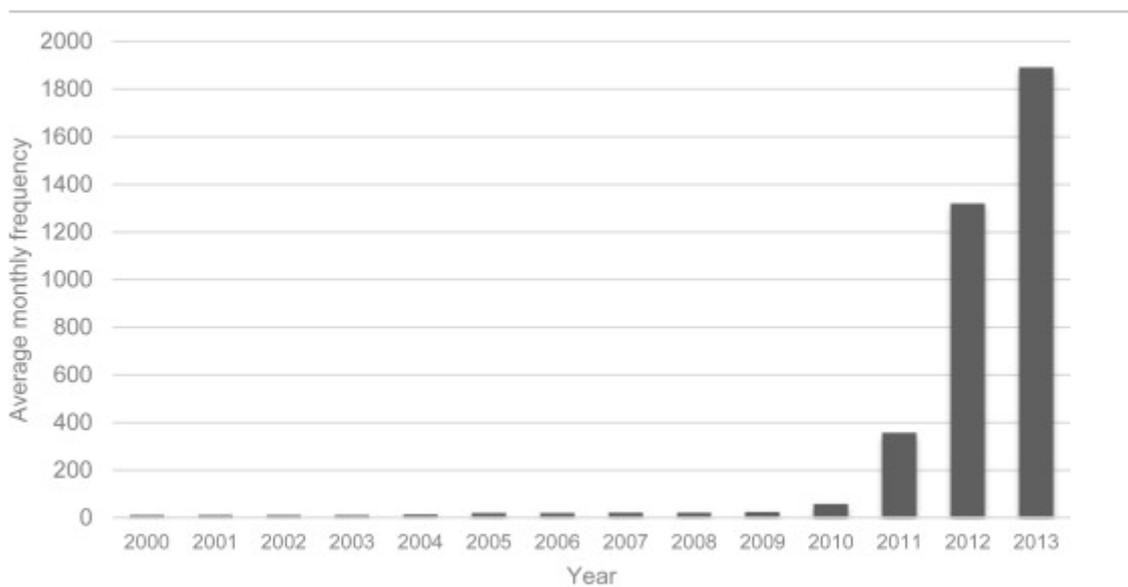


Figura 24: Distribuzione della frequenza del termine "Big Data" nella Research Library di ProQuest

Le definizioni di **big data** si sono evolute rapidamente, ciò ha portato grande confusione attorno al termine. Questa evidenza è stata dimostrata da un sondaggio online condotto da Harris Interactive per conto di SAP su 154 dirigenti nell'aprile 2012 (*"Small and mid-size companies look to make big gains with big data"*, 2012). La Fig. 25 mostra come i dirigenti differissero nella loro comprensione dei big data: alcune definizioni si concentravano su cosa fossero, mentre altre cercavano di rispondere a cosa fanno.

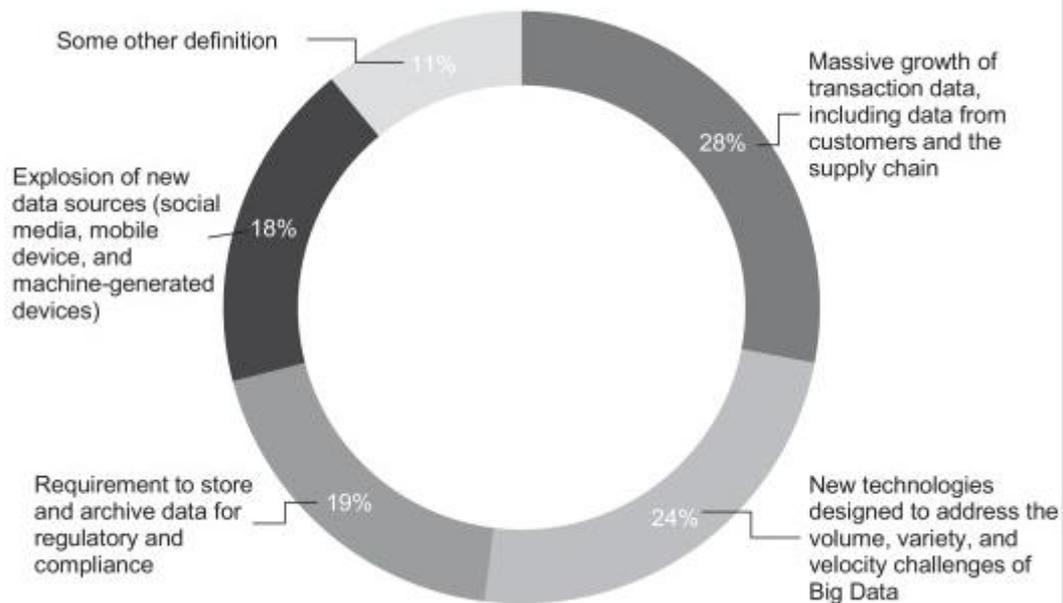


Figura 25: Definizioni di "big data" derivanti da un sondaggio online dell'aprile 2012 somministrato a 154 dirigenti

La dimensione è sicuramente la prima caratteristica che viene in mente, tuttavia, di recente sono emerse altre caratteristiche dei big data. Laney (2001) suggerì che Volume, Varietà e Velocità (o le Tre V) sono le tre dimensioni principali nella sfida alla gestione dati. Le Tre V sono emerse come un quadro comune per descrivere i big data (Chen et al., 2012, Kwon et al., 2014).

Gartner, Inc. definisce i big data in termini simili ("Gartner IT Glossary, n.d."):

"I big data sono risorse informative ad alto volume, ad alta velocità e di grande varietà che richiedono forme di elaborazione delle informazioni economiche e innovative per una migliore comprensione e per prendere decisioni".

Allo stesso modo, TechAmerica Foundation definisce i big data come segue (TechAmerica Foundation's Federal Big Data Commission, 2012):

"Big data è un termine che descrive grandi volumi di dati ad alta velocità, complessi e variabili che richiedono tecniche e tecnologie avanzate per consentire l'acquisizione, l'archiviazione, la distribuzione, la gestione e l'analisi delle informazioni".

Riprendendo queste definizioni si osserva come i dati siano diventati il nuovo carburante delle grandi aziende e PMI, e la vera chiave del successo è saper estrarre valore da questi.

- 1. Volumi:** la quantità dei dati, sia che essi siano stati generati da utenti sia che, in maniera automatica, da macchine. Possiamo ricondurre il volume dei dati generati da una azienda all'ordine di grandezza di terabyte o petabyte. Tuttavia, le definizioni dei volumi di big data sono relative e variano in base a fattori, come il tempo e il tipo di dati. Ciò che oggi può essere considerato big data potrebbe non raggiungere la soglia in futuro perché le capacità di archiviazione aumenteranno, consentendo l'acquisizione di set di dati ancora più grandi. Inoltre, il tipo di dati, esposto qui sotto alla voce "varietà", definisce cosa si intende per "grande". Due set di dati della stessa dimensione possono richiedere diverse tecnologie di gestione dei dati in base al loro tipo, ad esempio dati tabulari rispetto a dati video. Queste considerazioni rendono quindi poco pratico definire una soglia specifica per i volumi di big data.
- 2. Varietà:** si riferisce all'eterogeneità strutturale in un dataset. I progressi tecnologici consentono alle aziende di utilizzare vari tipi di dati strutturati, semi-strutturati e non strutturati. I dati strutturati, che costituiscono solo il 5% di tutti i dati esistenti (*Cukier, 2010*), si riferiscono ai dati tabulari trovati in fogli di calcolo o database relazionali. Testo, immagini, audio e video sono esempi di dati non strutturati, che a volte mancano dell'organizzazione strutturale richiesta dalle macchine per l'analisi. In un continuum tra dati completamente strutturati e non strutturati, il formato dei dati semi-strutturati non è conforme a standard rigorosi, sono informazioni che non risiedono in un database relazionale o in qualsiasi altra tabella di dati, ma presentano tuttavia alcune proprietà organizzative per semplificare l'analisi, come per esempio tag semantici. Un alto livello di varietà, caratteristica distintiva dei big data, non è necessariamente una novità per l'azienda. Le organizzazioni hanno sempre accumulato dati non strutturati da fonti interne (ad esempio: dati dei sensori) e fonti esterne (ad esempio: Social media). Tuttavia, l'aspetto innovativo è l'emergere di nuove tecnologie di gestione e analisi dei dati, che consentono alle organizzazioni di sfruttare questi dati nei loro processi aziendali.
- 3. Velocità:** si riferisce alla velocità con cui vengono generati i dati e alla necessità che questi arrivino in tempo reale al fine di effettuare analisi su di essi. La diffusione di dispositivi digitali come smartphone e sensori ha portato a un tasso di creazione di dati senza precedenti e sta generando un crescente bisogno di analisi in tempo reale e pianificazione basata su evidenze. Anche i retailers convenzionali generano dati molto rapidamente. Wal-Mart, ad esempio, elabora più di un milione di transazioni all'ora (*Cukier, 2010*). I dati provenienti dai dispositivi mobili e che fluiscono attraverso le app producono enormi quantità di informazioni che possono essere utilizzate per generare offerte personalizzate in tempo reale per i clienti. Questi dati forniscono informazioni valide sui clienti, come la posizione geografica, dati demografici e modelli di acquisto passati, che possono essere analizzati in tempo reale per creare valore reale per il cliente. I tradizionali sistemi di gestione dei dati non sono in grado di gestire enormi flussi di informazioni istantaneamente. È qui che entrano in gioco le tecnologie dei big data, strumenti in grado di garantirne il corretto immagazzinamento e l'efficiente analisi.

La 4ª e la 5ª V

Con il tempo, sono state introdotte una quarta V, quella di veridicità, e poi una quinta, quella di Valore:

4. **Veridicità:** rappresenta la qualità e l'inaffidabilità insita in alcune fonti di dati. Ad esempio, i sentimenti dei clienti nei social media sono di natura incerta, poiché implicano il giudizio umano. Tuttavia, contengono informazioni preziose. Pertanto, la necessità di gestire dati imprecisi e incerti è un altro aspetto dei big data, che viene affrontato utilizzando strumenti e analisi sviluppati per la gestione e l'estrazione di dati incerti.
5. **Variabilità:** SAS ha introdotto la variabilità come dimensione aggiuntiva dei big data. La variabilità si riferisce alla variazione delle portate di dati e al significato che tali informazioni hanno in base al contesto in cui vengono generate. Ciò impone una sfida critica: la necessità di connettere, abbinare, ripulire e trasformare i dati ricevuti da diverse fonti.



Figura 26: Le 5 V dei Big Data

Come possono essere utilizzati i Big Data

L'importanza dei Big Data va oltre la dimensione puramente numerica. Evolvono le metodologie di analisi dati, le infrastrutture e le tecnologie. I dati stessi assumono un nuovo ruolo nell'ecosistema economico e sociale e la figura professionale del Data Scientist cresce di rilevanza e competenze.

Oggi non approcciarsi ai *Big Data Analytics* equivale a perdere opportunità, a essere esclusi da nuovi mercati o addirittura da quelli in cui già si opera.

Sono diversi gli obiettivi aziendali raggiungibili grazie ai Big Data. Le imprese che hanno già messo in atto progetti di *data analysis*, *business intelligence* e *data management* lo hanno fatto per i seguenti obiettivi:

- migliorare l'engagement con il cliente
- incrementare le vendite
- ridurre il time to market
- ampliare l'offerta di nuovi prodotti e servizi
- ottimizzare l'offerta attuale al fine di aumentare i margini
- ridurre i costi
- identificare nuovi mercati

I benefici aziendali e sociali dei Big Data

A oggi non vi sono testimonianze, tra coloro che hanno avviato delle progettualità, che dimostrino il fallimento dei Big Data. L'analisi dei dati, seppur più complessa, ha aperto scenari ed opportunità inesplorate, creando un enorme vantaggio competitivo.

I vantaggi dei Big Data Analytics sono trasversali e coinvolgono tutte le funzioni aziendali e tutti i processi, generando benefici quantificabili e non.

Ma i Big Data portano anche benefici sociali che vanno molto oltre fatturati, profitti e strategie di marketing aziendali. Da prerogativa del business, i Big Data diventano un patrimonio per la collettività, è così che i dati assumono un importante ruolo sociale perché in grado di creare nuovi posti di lavoro, ripristinare il budget delle Pubbliche Amministrazioni, ottimizzare i flussi turistici di un territorio e perfino salvare delle vite. Pur con tutti i distinguo del caso, non è cosa nuova l'utilizzo degli Analytics nel mondo sanitario o nella gestione dei disastri naturali.

Data Integration

Con il termine *Data Integration* si intendono tutte quelle pratiche, tecniche e strumenti che permettono di fornire un accesso unico e la messa a disposizione di tutti i dati necessari ai requisiti dei vari sistemi informativi aziendali nonché alle richieste dei processi di business²².

Riuscire ad integrare il maggior numero di dati possibili è fondamentale per diversi motivi, in particolare per ricercare pattern e relazioni tra diverse fonti oppure per gestire tipologie di dati, quali dati semi-strutturati o non strutturati, non immagazzinabili attraverso tecnologie tradizionali.

Come già detto in precedenza, il valore che si può estrarre dai Big data non proviene tanto dalla mole di dati, quanto invece dalla correlazione di una varietà di fonti, tipologie e formati di dati.

Tuttavia, a causa della mancanza di tecnologie adeguate e l'integrazione e la governance di dati eterogenei è una sfida che molte imprese non riescono ad affrontare in maniera ottimale.

In seguito, verranno mostrate le possibili architetture per l'immagazzinamento dei Big Data.

Il livello di integrazione dei dati: quattro approcci

Possiamo sintetizzare i livelli di integrazione attraverso quattro approcci, di seguito elencati in ordine crescente in base al livello di integrazione:

- Silos
- Data Warehouse (DWH),
- Data Lake (DL),
- Modello Integrato, modello secondo il quale l'organizzazione ha a disposizione sia un Data Lake che un Data Warehouse che lavorano in modalità integrata per rispondere alle differenti esigenze di storage, gestione e analisi di qualsiasi tipologia di dato.

²² <https://www.gartner.com/en/information-technology/glossary/data-integration-tools>

Data Warehouse

Un data warehouse è un'architettura di archiviazione progettata per contenere dati estratti da sistemi transazionali, archivi operativi e in generale sorgenti esterne. Il DWH combina questi dati (per esempio aggregandoli) in modo che siano disponibili in una forma adatta per effettuare l'analisi dei dati e creazione di report che soddisfano le richieste del business.

La struttura del database viene definita a priori (Schema-on-Write²³), i dati vengono scritti in questa struttura e quindi analizzati.

Le sue caratteristiche e modalità di accesso sono perfette per utenti che hanno bisogno di analisi abbastanza semplici e, soprattutto, che sanno cosa cercare, quale tipo di relazione analizzare. Sono utenti di business che trovano in questi "contenitori" e nei sistemi di business intelligence che li interrogano ottime risposte in termini di reportistica e per capire l'andamento di fenomeni conosciuti.

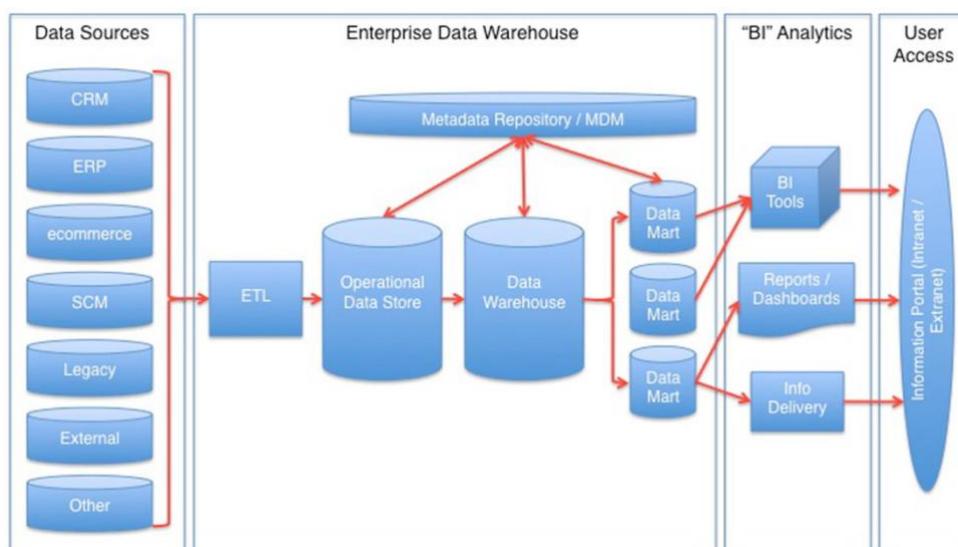


Figura 27: Architettura di riferimento di un data warehouse

Data Lake

Luogo di archiviazione in grado di ospitare dati grezzi (*raw data*) fin quando non è necessario dar loro una struttura. Con questa modalità di gestione è possibile avere l'integrazione di elevate quantità di dati di qualsiasi formato e provenienti da qualsiasi fonte.

I dati raggiungono il luogo d'archiviazione e vengono esaminati e campionati nei loro formati originali: solo nel momento in cui vengono interrogati da applicazioni e operatori vengono convertiti in formati leggibili dai sistemi aziendali e quindi possono essere confrontati con altre informazioni.

I *Data Lake* possono contenere dati strutturati, semi-strutturati (CSV, log, XML, JSON) e non strutturati, ogni elemento viene identificato in modo univoco attraverso una serie di tag che

²³ Lo schema/struttura è definito in anticipo e poi i dati sono scritti in tale schema come parte del processo di scrittura.

corrispondono ai suoi metadati; ciò consente, nel momento in cui un *Data Lake* viene interrogato relativamente a un problema specifico, di estrarre i dati rilevanti e di analizzarli. Questa metodologia viene anche definita “Schema-on-Read²⁴” dove i dati vengono letti senza dover essere “scritti” secondo uno specifico modello.

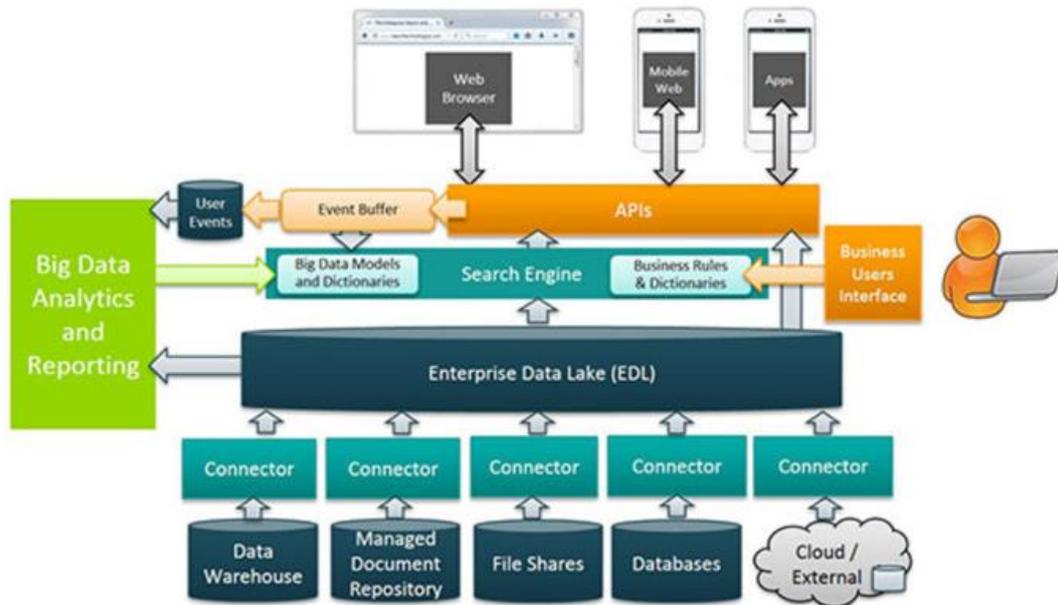


Figura 28: Esempio di architettura di un Data Lake

Il termine *Data Warehouse* e *Data Lake* si riferiscono dunque a due diverse strategie di archiviazione e gestione dei dati, dove le principali differenze riguardano la struttura, lo scopo a cui serviranno, gli utenti che andranno a utilizzarli e l’accessibilità dei dati raccolti.

Data Analytics

Data Analytics è una disciplina in grado di trasformare i dati grezzi in *insights* o modelli significativi utili per prendere decisioni migliori. Il termine *Data Analytics* comprende la raccolta dei dati, l’organizzazione, l’immagazzinamento, l’analisi e tutti gli strumenti e le tecniche utilizzate.

Possiamo raggruppare i modelli di *Analytics* in quattro macro-categorie:

- *Descriptive Analytics*, l’insieme di strumenti orientati a descrivere la situazione passata ed attuale dei processi aziendali e/o aree funzionali.
- *Predictive Analytics*, strumenti avanzati che attraverso l’utilizzo di tecniche matematiche — come ad esempio regressioni, forecasting e modelli predittivi — analizzano le informazioni e provano a rispondere a domande relative a cosa potrebbe accadere nel futuro.
- *Prescriptive Analytics*, tool in grado di proporre al *decision maker* soluzioni operative/strategiche sulla base delle analisi svolte.
- *Automated Analytics*, strumenti in grado di implementare autonomamente l’azione proposta secondo il risultato delle analisi svolte.

²⁴ I dati grezzi vengono prima inclusi nella loro forma nativa (strutturata e/o non strutturata) senza uno schema imposto. Solo dopo che i dati sono stati letti viene applicato lo schema.

La rilevanza dei Big Data per i Manager e i CIO

Il mercato degli Analytics non conosce crisi e si attesta a un valore complessivo di 1.4 miliardi di euro. Secondo un'indagine dell'Osservatorio Big Data Analytics & Business Intelligence, la gestione dei Big Data è ai vertici tra le priorità di investimento di CIO e Innovation Manager italiani.

L'interesse verso il Big Data Analytics e le scelte d'investimento sono differenti a seconda della dimensione delle imprese da analizzare. Per quanto riguarda grandi aziende, sono due le ottiche prevalenti. Da una parte le imprese innovative, che hanno già avviato progetti, si trovano a raccogliere i frutti dei primi investimenti e spingono ora su tecnologie sempre più evolute.

Dall'altro, le aziende "ritardatarie" corrono ai ripari, rinnovando processi e architetture in ottica Big Data. Discorso diverso per le PMI che, complice un ritardo sia culturale che tecnologico, procedono piuttosto a rilento.

Nell'elaborato non si sono trattati i *Data Lake* come possibili soluzioni ai *Data Silos* in quanto, a seguito di un'analisi preliminare sulle informazioni in esame, è subito emerso come queste non rispondessero, o lo facessero solo in parte, ai criteri inclusi nella definizione di *Big Data*.

Bibliografia e Sitografia

Feldman S., Sherman C. The High Cost of Not Finding Information. IDC White Paper.
<https://forcam.com/en/what-is-traceability>
<https://www.grantadesign.com/download/pdf/Granta-Business-Case-Italian.pdf>
<https://www.zerounoweb.it/software/cose-lo-sviluppo-low-code-e-quali-sono-le-piattaforme-disponibili>
<https://www.outsystems.com/blog/posts/benefits-of-low-code-platforms>
<https://www.business.org/finance/inventory-management/rfid-vs-barcodes-vs-qr-codes/#RFID>
<https://www.pmi.it/impresa/business-e-project-management/articolo/1755/agile-project-management.html>
<https://pacificbiolabs.com/pharmaceutical-cmc-expectations>
Takeuchi H, Nonaka I. 16 The new new product development game. Japanese Business: Part 1, Classics Part 2, Japanese management Vol.2: Part 1, Manufacturing and production Part 2, Automotive industry Vol.3: Part 1, Banking and finance Part 2, Corporate strategy and inter-organizational relationships Vol.4: 1998;64(1):321.
<https://agilemanifesto.org/iso/it/manifesto.html>
Annual State of Agile Report, VersionOne; 2010
<https://en.wikipedia.org/>
https://www.keyence.it/ss/products/marketing/traceability/basic_about.jsp
Augustine S. Managing agile projects: Prentice Hall PTR; 2005.
Manifesto for Agile Software Development: Kent Beck, Robert C. Martin, Martin Fowler; 2001
https://en.wikipedia.org/wiki/Information_silo
<https://maxxphase.com/data-compatibility/top-benefits/data-silo-elimination/>
<https://web.archive.org/web/20161005230416/https://www.rotize.com/silo-mentality/>
<https://www.ame.org/sites/default/files/documents/88q1a3.pdf>
<https://www.ame.org/sites/default/files/documents/89Q4A5.pdf>
<http://itre.cis.upenn.edu/~myl/language-log/archives/002964.html>
<https://web.archive.org/web/20120601005717/http://www.camagazine.com/archives/print-edition/2002/march/columns/camagazine23400.aspx>
<https://blog.hubspot.com/service/data-silos>
<https://www.cmswire.com/customer-experience/the-path-to-customer-centricity-lies-in-dismantling-data-silos/>
<https://humansofdata.atlan.com/2020/02/what-are-data-silos/>
https://www.google.com/search?q=vantaggio+big+data&rlz=1C1GCEA_enIT912IT912&oq=vantaggio+big+data&aqs=chrome..69i57j0.4203j1j7&sourceid=chrome&ie=UTF-8
https://blog.osservatori.net/it_it/big-data-a-cosa-servono#:~:text=L'analisi%20dei%20dati%2C%20pur,generando%20benefici%20quantificabili%20e%20non.
https://blog.osservatori.net/it_it/big-data-cosa-sono
https://blog.osservatori.net/it_it/data-integration-cosa-significa-come-farla?hsCtaTracking=9e8bf75d-f4a6-4591-9165-af24c5595fda%7Cfaa2fdaa-359e-4330-b5d1-97cfa1e01f83
https://blog.osservatori.net/it_it/data-lake-significato-vantaggi
https://blog.osservatori.net/it_it/big-data-a-cosa-servono
<https://searchaws.techtarget.com/definition/data-lake>
<https://www.spindex.it/it/blog/data-silos-data-lakes/>
<https://www.sciencedirect.com/science/article/pii/S0268401214001066>
<https://www.forbes.com/sites/gartnergroup/2013/03/27/gartners-big-data-definition-consists-of-three-parts-not-to-be-confused-with-three-vs/#32f9d17742f6>
<https://gurukul.com/blog/big-data-analytics-ueba-data-lake-warehouse>

<https://www.gartner.com/en>

<https://www.zerounoweb.it/analytics/big-data/big-data-gestirli-in-4-mosse/>

<https://www.zerounoweb.it/analytics/analytics-cosa-significa-quando-e-come-si-usa/>

<https://searchcustomerexperience.techtarget.com/definition/citizen-development>

Ringraziamenti

Un ringraziamento a tutti i professori per avermi permesso di crescere non solo come studente ma soprattutto come persona, grazie in particolare alla professoressa Alessandra Colombelli per la disponibilità che mi ha concesso, per la professionalità che ha dimostrato e per il supporto dato in questo momento cruciale della mia carriera universitaria

All'azienda Chiesi Farmaceutici per l'esperienza offertami e al mio tutor Roberto Gaspari che nonostante gli innumerevoli impegni, si è sempre reso disponibile nell'aiutarmi

Ai miei genitori, per i sacrifici fatti e per avermi permesso di essere qui, oggi, a festeggiare

A Irene che nonostante i litigi mi è sempre stata vicino cercando nei momenti più tesi di strapparmi un sorriso

A Rita per avermi dispensato consigli nelle difficoltà, ad Alessandro mio più grande punto di riferimento e a cui devo la persona che sono diventato, al piccolo Leonardo che con ogni suo sorriso ha illuminato le giornate più grigie. A loro che nonostante la distanza non hanno mai fatto mancare il loro sostegno

Alla mia seconda famiglia, grazie a Stefano, Lorena, Marta, Luca e Lucia per tutto quello che fanno e hanno fatto per me

Ai parenti, ai ragazzi del bar, agli amici che ho conosciuto in questa bellissima esperienza e a quelli che ci sono sempre stati. Grazie per aver condiviso con me le esperienze più importanti, vi voglio bene.

Alla persona più speciale, Matilde, per avermi supportato e sopportato, perché la sua presenza è stata sempre la mia forza. Grazie per aver festeggiato con me ogni successo e per avermi aiutato a rialzarmi dopo ogni caduta, ma oggi, anche oggi, abbiamo di nuovo vinto noi.

Infine, grazie a me...perché è giusto così