



**UNIVERSITÀ POLITECNICA DELLE MARCHE
FACOLTÀ DI ECONOMIA “G. FUÀ”**

Corso di Laurea Magistrale in Economia e Management
Curriculum Amministrazione, Finanza e Controllo

**CRM E BUSINESS INTELLIGENCE, UN CONNUBIO
IDEALE PER L'ANALISI: IL CASO DEI LASCITI TESTAMENTARI ALLA
LEGA DEL FILO D'ORO ONLUS**

**CRM AND BUSINESS INTELLIGENCE, AN IDEAL COMBINATION
FOR ANALYSIS: THE CASE OF TESTAMENTARY BEQUESTS TO
LEGA DEL FILO D'ORO ONLUS**

Relatore:
Chiar.mo Prof. Danilo Scarponi

Tesi di Laurea di:
Veronica Gaggiano

Anno Accademico 2018/2019

INDICE

INTRODUZIONE.....	6
CAPITOLO 1: KNOWLEDGE MANAGEMENT, SISTEMI INFORMATIVI AZIENDALI E CRM.....	9
1.1 LA GESTIONE DEI DATI IN DATABASE	9
1.2 LA GESTIONE DELL'INFORMAZIONE	11
1.3 I SISTEMI INFORMATIVI AZIENDALI	13
1.3.1 Le caratteristiche dei sistemi informativi	16
1.4 IL CRM	20
CAPITOLO 2: IL DATA WAREHOUSE.....	29
PREMESSA	29
2.1 COS'È UN DATA WAREHOUSE.....	29
2.2 ELEMENTI COSTITUTIVI DI UN DATA WAREHOUSE.....	31
2.3 I DATA MART	33
2.4 ARCHITETTURE DI DATA WAREHOUSE.....	37
2.5 QUALITÀ DEI DATI	46
2.6 GLI STRUMENTI DI ETL	50
2.7 OLAP.....	51
CAPITOLO 3: IL DATA MINING.....	55
3.1 IN COSA CONSISTE IL DATA MINING	55
3.2 CREAZIONE DI UN SISTEMA DI DATA MINING	57
3.3 LE FUNZIONI DI MINING.....	59

3.4 METODOLOGIE DI DATA MINING.....	59
3.5 GLI STRUMENTI DI MINING DEI DATI	63
CAPITOLO 4: LA CONOSCENZA DEL “SOGGETTO BENEFICIANTE”: IL CASO “LEGA DEL FILO D’ORO ONLUS”	68
PREMESSA	68
4.1 LA STORIA DELLA LEGA DEL FILO D’ORO.....	69
4.2 L’IMPORTANZA DELLA RACCOLTA FONDI	74
4.3 IL TESTAMENTO SOLIDALE	77
4.4 L’INTRODUZIONE DEL CRM	81
4.5 IL PROGETTO “UN RICORDO TANGIBILE”	84
CONCLUSIONI	103
BIBLIOGRAFIA	106
SITOGRAFIA	108

INTRODUZIONE

La globalizzazione, la crisi economica mondiale, la sempre maggiore competitività, unite all'innovazione e alla continua ricerca dell'elevata qualità hanno portato le imprese a cercare di fare sempre meglio, di soddisfare al massimo i clienti ed ottenere così un vantaggio competitivo.

È possibile ottenere tali vantaggi in vari modi. Un elemento divenuto particolarmente strategico in questo senso è la conoscenza, un asset intangibile ormai divenuto di estrema importanza per perseguire gli obiettivi di cui sopra.

Oggi infatti si parla sempre più di "knowledge management" per indicare questa continua ricerca di informazioni che dovranno essere prodotte a partire dai dati raccolti dai sistemi informativi aziendali interni all'azienda e che dovranno essere appunto gestiti. Ciò ha permesso di prendere in considerazione una serie di aspetti innovativi mai considerati in precedenza, in modo da raccogliere una mole di dati e trasformarli in informazioni di carattere gestionale e strategico sui prodotti, clienti, processi aziendali.

In particolare, si è passati dalla focalizzazione sul prodotto ad una visione "cliente-centrica" ossia basata sulla relazione con il cliente per cercare di soddisfare al meglio i propri bisogni ed anzi cercare di anticiparli.

L'obiettivo diventa perciò quello di studiare il cliente, le sue abitudini di acquisto, le preferenze nella comunicazione, puntando al soddisfacimento dei bisogni, seppur specifici, e cercare soprattutto di anticiparli per costruire relazioni di lungo periodo.

Tramite i sistemi informativi aziendali vengono raccolte una serie di informazioni utili sia nel presente, sia in vista di ulteriori analisi per il futuro. Si tratta di dati che inizialmente potrebbero anche non essere utili perché mai analizzati da nessuno prima e proprio per questo potrebbero diventare fondamentali per anticipare le mosse dei concorrenti.

Dai sistemi info aziendali si procede all'estrazione dei dati sotto forma di tabelle multidimensionali, che permettono di avere a disposizione tutti i dati selezionati in base agli obiettivi di analisi, creando il datawarehouse.

Attraverso gli strumenti di data mining che sono alla base dei sistemi di business intelligence, sarà possibile dedurre una serie di informazioni a partire dal datawarehouse che verrà poi ulteriormente scomposto in data mart ossia in un sottosistema tematico di datawarehouse utile ad effettuare un'analisi più puntuale.

Tali informazioni saranno estremamente utili al processo decisionale, se ben selezionate e rielaborate, e rivestiranno importanza fondamentale ai fini dell'ottenimento del vantaggio competitivo.

Lo scopo del presente elaborato è quello di offrire spunti di riflessione sull'importanza della conoscenza e su come questa può essere gestita, ossia come i dati contribuiscono a produrla ed a renderla fruibile in virtù di una futura crescita dell'impresa.

In questo senso si sottolinea l'importanza della creazione di un sistema informativo adeguato all'obiettivo da raggiungere e quali potenzialità questo può offrire.

Grazie alle informazioni inserite nel CRM utilizzato presso la Lega del Filo d'Oro Onlus si è potuto creare uno schema di data warehouse andando ad effettuare operazioni di miglioramento della qualità dei dati inseriti anche attraverso la redazione di una procedura, sulla base del funzionamento del CRM e dell'esperienza effettuata, in accordo con la Direzione.

Creato il data warehouse si sono stilate alcune analisi con Microsoft Power BI (versione desktop) una soluzione di Business Intelligence che permette di effettuare analisi dei dati da presentare tramite dashboard e schemi interattivi che permettono la fruibilità e la navigabilità.

Lo scopo di questo elaborato è uno studio sulla struttura ed il funzionamento dei dati esistenti con l'obiettivo di ottenere informazioni utili ad effettuare previsioni e iniziative mirate in vista degli obiettivi futuri da perseguire.

CAPITOLO 1: KNOWLEDGE MANAGEMENT, SISTEMI

INFORMATIVI AZIENDALI E CRM

1.1 LA GESTIONE DEI DATI IN DATABASE

La gestione dei dati in un unico sistema è conseguenza del fatto che in passato gli stessi dati, se presenti in file diversi, potevano portare ad uno spreco di tempo e di risorse ma anche ridondanze, confusione, scarsa sicurezza nei processi decisionali, e problemi di modifica dello stesso dato.

Uno dei principali problemi dell'organizzazione tradizionale dei dati è la **ridondanza** ossia "la presenza di dati duplicati in diversi file, che fa sì che gli stessi dati siano memorizzati in più di un luogo. La ridondanza dei dati si verifica quando vi sono divisioni, aree funzionali, e gruppi differenti all'interno di un'azienda che raccolgono indipendentemente gli stessi dati e li memorizzano in modo reciprocamente indipendente"¹. Per questi motivi sono nati i database che permettono agli utenti delle diverse aree aziendali di operare sugli stessi dati in un unico "luogo" come lo definiscono gli autori, non inteso in senso fisico ma logico.

¹ Laudon K. C., Laudon J. P., Morabito V., Pennarola F., Management dei sistemi informativi vol. 1 Fondamenti, Terza edizione, Pearson Prentice Hall, 2010.

Su questa via, le aziende hanno iniziato a utilizzare il cosiddetto DBMS (Database Management System) “un software che consente all'azienda di centralizzare i dati, gestirli in modo efficiente e fornire l'accesso ai dati dai programmi applicativi”² in maniera semplice e immediata.

L'attenzione nella creazione di un database deve essere infatti rivolta alle relazioni tra i dati (esistenti o eventualmente da creare) e una volta individuate, si dovrà definire il modello concettuale-logico del datawarehouse, costituito da tabelle di dati che saranno collegate tra loro da relazioni che possono essere di tipo “uno a uno” oppure “uno a molti” o ancora “molti a uno”.

Molto utile risulterebbe l'adozione di una *Politica di gestione dei dati e delle informazioni*, per stabilire chi può accedere ai dati e chi può modificarli visto che sempre più spesso si sente parlare di dati sensibili e della protezione di questi ultimi. Infatti, i responsabili dei dati dovrebbero garantire la qualità dei dati che analizzano, la sicurezza, la conformità alle leggi e l'accuratezza utile a produrre le informazioni di cui si ha bisogno. Ovviamente queste politiche saranno differenti anche a seconda delle dimensioni aziendali e del tipo di attività svolta.

² Laudon K. C., Laudon J. P., Morabito V., Pennarola F., Management dei sistemi informativi vol. 1 Fondamenti, Terza edizione, Pearson Prentice Hall, 2010, pag. 83.

1.2 LA GESTIONE DELL'INFORMAZIONE

I dati raccolti secondo logiche prestabilite, una volta messi insieme, saranno analizzati e rielaborati in modo da produrre l'informazione utile ai decisori. Non è detto però che tutte le informazioni ottenute vengano prese in considerazione e per questo bisogna selezionarle, o meglio, gestirle.

Sempre più spesso si sente parlare dell'importante ruolo che rivestono i sistemi informativi nella cosiddetta “**Knowledge Management**” ed è proprio la gestione della conoscenza l'elemento che permetterebbe alle imprese di diventare più efficienti ed ottenere un vantaggio competitivo rispetto ai concorrenti. La conoscenza, sotto forma di informazioni disponibili può essere considerata come un asset fondamentale, un “bene strategico la cui corretta gestione può garantire la sopravvivenza stessa di un'organizzazione o la differenziazione rispetto agli altri operatori del proprio settore”³.

Negli ultimi anni si è assistito ad un forte cambiamento dell'ambiente in cui le imprese operano, ambiente sempre più competitivo ed esposto ai rischi derivanti

³ TAGLIAVINI M., RAVARINI A., SCIUTO D., *Sistemi per la gestione dell'informazione*, Apogeo, Milano, 2003, pag. 2/3.

dall'apertura al mercato globale e dalle forti spinte tecnologiche (Inghirami parla di “ipercompetizione”⁴). Questi elementi hanno portato ad un cambiamento dei consumi e degli stili di vita e perciò alla continua ricerca e individuazione di nuovi fattori critici di successo da parte delle imprese.

In un mercato di questo tipo, in cui si parla anche di mercato “individualizzato”⁵ (A. De Luca, 2011) riferito cioè alla soddisfazione mirata del singolo cliente, i fattori chiave per ottenere successo e battere i concorrenti sono costituiti dall'informazione e dalla conoscenza, e non più dal prodotto. Per questo motivo si è diffusa negli ultimi anni una nuova filosofia di gestione denominata “Knowledge Management” (Gestione della Conoscenza), considerata ormai elemento irrinunciabile per ottenere successo.

Per far ciò le imprese devono dotarsi di un efficace sistema informativo in grado di raccogliere dati e informazioni utili sia ai fini conoscitivi (per analizzare ciò che sta accadendo) sia ai fini decisionali (per capire cosa potrebbe accadere).

La gestione dell'informazione riguarderà una serie di attività tra cui:

- *Creare* l'informazione, cioè introdurre qualcosa che prima non esisteva;
- *Acquisire* l'informazione dai processi aziendali;

⁴ Cfr. I. E. Inghirami, Il sistema informativo amministrativo per la governance aziendale, Franco Angeli, 2008, pag. 21

⁵ De Luca A., Amministrazione e finanza nelle PMI innovative. Nuove metriche per valutare la performance, Franco Angeli, Milano, 2011.

- *Elaborare* l'informazione, leggendo una serie di dati si traggono informazioni complessive riferite ad un certo fenomeno;
- *Archiviare* l'informazione, cioè conservare le informazioni affinché gli sia assicurata una longevità nel tempo;
- *Trasmettere* l'informazione ai soggetti interessati;
- *Presentare* l'informazione⁶.

Come si nota, le attività da svolgere sono diverse e molto specifiche, perciò richiedono molto tempo, risorse finanziarie elevate e personale dedicato.

1.3 I SISTEMI INFORMATIVI AZIENDALI

La centralità delle informazioni, insieme allo sviluppo dell'informatica, ha portato già da molti anni all'introduzione di sistemi in grado di raccogliere dati per poterli poi rielaborare ed utilizzare per i fini decisionali aziendali. Questi sistemi definiti "Sistemi informativi" rappresentano strumenti di raccolta dei dati di (potenziale) interesse per l'azienda derivanti dalle vicende aziendali interne ed esterne, per poterli poi rielaborare, trasformandoli in informazioni fruibili da parte dei responsabili di funzione e soprattutto dai vertici aziendali.

⁶ TAGLIAVINI M., RAVARINI A., SCIUTO D., *Sistemi per la gestione dell'informazione*, Apogeo, Milano, 2003, pag. 3/4.

Diverse sono le definizioni di “Sistemi informativi” riscontrabili in letteratura, per cui non esiste una definizione univoca.

Alcuni autori si focalizzano sulla caratteristica di essere un *software* di raccolta delle informazioni mentre altri si concentrano piuttosto sulla loro finalità in azienda.

Una definizione abbastanza completa, a mio avviso, è quella di Pighin e Marzona secondo cui “i sistemi informativi aziendali hanno come obiettivo la guida nell’esecuzione dei processi aziendali e la presentazione di informazioni alle persone nel momento in cui tali informazioni sono necessarie. Questi compiti sono assolti tramite la definizione di procedure che permettono la raccolta dei dati in archivi organizzati, l’elaborazione degli stessi fino al loro consolidamento in informazioni e la distribuzione di queste agli utenti”⁷. Infatti, tramite l’utilizzo di tali sistemi è possibile ricercare in ogni momento e in maniera molto rapida qualunque tipo di informazione sul cliente, dai dati anagrafici, alle richieste di informazioni effettuate in passato, i soggetti aziendali con i quali ha avuto il contatto, fino agli indicatori generali di performance (KPI: Key Performance Indicator) che riguardano ad esempio il numero delle vendite, la redditività, la durata della relazione e così via.

⁷ PIGHIN M., MARZONA A., *Sistemi informativi aziendali, ERP e sistemi di data analysis*, Terza edizione, Pearson, 2018, pag. 4-6.

Ancora più importate in questa definizione è il riferimento alle procedure di inserimento dei dati, necessarie per ottenere dati standardizzati e di buona qualità in modo da rendere più agevole e precisa l'analisi, oltre che la loro interpretazione. Secondo Candiotta (2004) invece, "attengono al sistema informativo tutte le attività di raccolta, elaborazione e trasmissione dei dati realizzate al fine di soddisfare le esigenze conoscitive interne ed esterne"⁸.

I dati riferiti alle vicende aziendali interne ed esterne rappresentano il punto di partenza per la produzione di informazioni, dato che questi dovranno essere decifrati e rielaborati per trasformarsi in informazioni utili a migliorare i processi interni e ad offrire migliori servizi verso l'esterno (clienti e stakeholders).

Sempre secondo l'autore, "il sistema informativo ha lo scopo di produrre e distribuire informazioni qualificate a tutti i soggetti che in vario modo se ne servono per svolgere le proprie funzioni, facilitandone il processo decisionale e lo svolgimento dei compiti"⁹. Sicuramente uno dei vantaggi che questi sistemi offrono è dato dal fatto che le informazioni sono fruibili da parte dei soli utenti interessati e che vengono aggiornate in tempo reale dagli utenti stessi, perciò eventuali modifiche o correzioni andranno a beneficio di tutti contemporaneamente.

⁸ Candiotta R., Sistemi informativi integrati, Giuffrè Editore, 2004, pag. 1.

⁹ Candiotta R., Sistemi informativi integrati, Giuffrè Editore, 2004, pag. 13.

1.3.1 Le caratteristiche dei sistemi informativi

Gli elementi costitutivi di un sistema informativo sono (Figura 6):

1. *L'insieme di dati riferiti alla realtà aziendale e all'ambiente esterno;*
2. *Il complesso delle procedure per la raccolta dei dati, l'elaborazione e distribuzione delle informazioni;*
3. *Le risorse umane e tecnologiche impiegate nell'attuazione del processo informativo.*

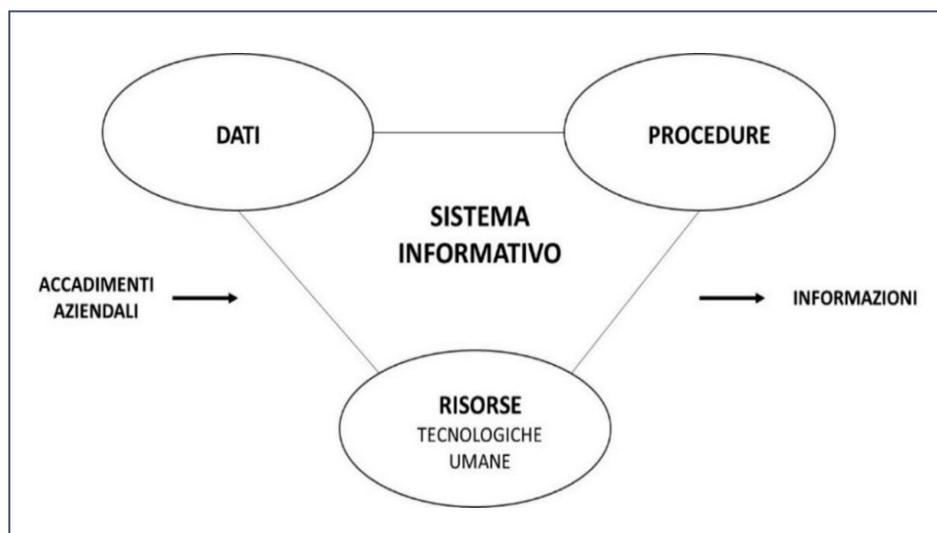


Figura 1 Gli elementi del sistema informativo aziendale (fonte R. Candiotta, opera citata)

Riguardo la tipologia, i sistemi informativi possono distinguersi in:

- *Formali*, quando contengono informazioni specifiche e che vengono rilevate seguendo procedure ben definite;
- *Informali*, quando le informazioni che contengono sono riferite a rapporti informali (telefonate, notizie, scambio di opinioni tra colleghi).

Molto importante in questi sistemi risulta l'attività di alimentazione e le relative modalità con le quali avviene tale attività. È fondamentale la selezione dei dati, la loro classificazione, la codificazione affinché producano informazioni rilevanti. Infatti dopo queste attività, lo step che segue è quello dell'elaborazione dei dati che prevede l'attività di verifica dei dati e la loro memorizzazione, che dovrà essere piuttosto standardizzata e soprattutto modificata in base alle esigenze informative dei destinatari¹⁰.

L'ultimo step è quello di trasmissione delle informazioni, in cui si va a comunicare all'utente il risultato finale del processo informativo.

Tagliavini et al. (2003 opera citata), parlando di sistemi informativi fanno riferimento oltre che agli strumenti hardware e software, anche a tutto l'insieme di attività ed elementi (persone, dati, processi, ...) che "interagiscono con le informazioni gestite in una data organizzazione". Proprio le persone ossia gli **utenti** dei diversi livelli organizzativi, costituiscono la **parte attiva** del sistema

¹⁰ Per ulteriori approfondimenti si rimanda a: Candiotta R., Sistemi informativi integrati, Giuffrè Editore, 2004, pagg. 9-11.

informativo in quanto svolgono le attività operative che producono dati, alimentano il sistema oppure usufruiscono dei dati per effettuare analisi di Business Intelligence. Sarebbe quindi errato (secondo gli autori) attribuire la responsabilità dei dati al soggetto cui fa capo la gestione delle tecnologie informatiche, che avrebbe l'esclusivo compito di manutenzione e aggiornamento dei componenti software e hardware.

Negli ultimi anni si è assistito al passaggio dal sistema informativo contabile al sistema informativo amministrativo. Alla base di questo cambiamento vi è lo spostamento dell'ottica di riferimento: da una visione "ex post" avente l'obiettivo di analizzare i dati dopo averli registrati per avere una visione d'insieme sugli avvenimenti passati, ad una "ex ante" volta a capire quale potrebbe essere l'evoluzione futura di alcuni avvenimenti per fare da supporto alla presa di decisioni da parte del management¹¹.

Da qui, la scomposizione dei **sistemi informativi** tra:

- **operazionali**: in cui vengono registrate quotidianamente le attività semplici e routinarie svolte in azienda e le transazioni con l'esterno;

¹¹ Cfr. MARCHI L., MANCINI D. (a cura di), *Gestione informatica dei dati aziendali*, Terza edizione, Franco Angeli, Milano, 2009.

- **direzionali:** raccolgono ed elaborano i dati utili principalmente nel lungo periodo per identificare i bisogni aziendali futuri e stabilire le strategie da mettere in atto.

L'utilità del sistema informativo è data principalmente dal fatto che l'impresa nello svolgimento delle proprie attività deve produrre informazioni a supporto delle decisioni. In effetti, "l'attività direttiva può essere intesa come il processo di conversione delle informazioni in decisioni"¹² e "la capacità direttiva dipende dal modo in cui sono utilizzate le informazioni disponibili al fine di razionalizzare l'assunzione di decisioni"¹³. Come afferma Bill Gates "Come raccogli, gestisci e utilizzi le informazioni determinerà la tua vittoria o la tua sconfitta": da qui si evince come le informazioni siano elemento imprescindibile di successo.

Come già ribadito, è grazie allo sviluppo dell'informatica che sono nati applicativi in grado di gestire un'enorme mole di dati che sarebbe quasi impossibile analizzare da parte di un essere umano. Il primo passo compiuto in questo senso ha portato alla creazione dei Data Warehouse, ossia basi di dati che hanno come scopo finale quello di produrre report che dovranno poi essere analizzati più in profondità

¹² MAURI C., *Il sistema informativo per il marketing. Dai dati alle risorse di conoscenza per la gestione delle relazioni con i clienti*, Giappichelli Editore, Torino, 2007.

¹³ MAURI C., *Il sistema informativo per il marketing. Dai dati alle risorse di conoscenza per la gestione delle relazioni con i clienti*, Giappichelli Editore, Torino, 2007 pag. 1.

attraverso l'utilizzo degli strumenti di data mining (aspetti che saranno approfonditi nei successivi capitoli).

1.4 IL CRM

Nell'ambito dei sistemi informativi aziendali assume importanza sempre più strategica il CRM (Customer Relationship Management)-

L'attuale contesto competitivo si caratterizza per un ciclo di vita dei prodotti sempre più breve e tecnologie in continuo avanzamento che comportano per le aziende la necessità di adeguamento continuo per ottenere dei vantaggi competitivi.

Il vantaggio competitivo sarà ottenuto da quelle aziende in grado non solo di distinguersi dalla concorrenza e rendersi "inimitabili", ma che sono soprattutto in grado di fidelizzare i propri clienti. Ed è proprio la fidelizzazione del cliente che si punta a conseguire ricorrendo all'utilizzo delle soluzioni di CRM che tengono traccia di tutto ciò che riguarda la relazione tra l'azienda e il cliente, a partire dalle prime fasi di contatto, passando poi a quelli di acquisizione del cliente stesso fino all'interruzione del rapporto. La relazione costituisce infatti oltre che una fonte di

vantaggio competitivo, un elemento intangibile determinante il valore d'impresa: si parla in questo senso di *capitale relazionale*.

Riportando le parole di Marchi e Mancini, "il Customer Relationship Management rappresenta la risposta a questa esigenza delle organizzazioni di migliorare la loro capacità di gestire le relazioni con i propri clienti "fidelizzati" ed accrescere il loro valore prospettico"¹⁴. L'obiettivo che le aziende dovranno perseguire sarà quello di aumentare la profittabilità che deriva dai clienti attraverso la fidelizzazione. Dopo aver analizzato le caratteristiche di ogni singolo cliente, si dovranno mettere in atto azioni di marketing che permettano di *consolidare* la relazione e di ottenere un profitto sempre maggiore. Nel caso in cui però il cliente non genera redditività e non risponde in maniera positiva alle campagne di fidelizzazione, genererà solo dei costi per l'azienda, la quale avrà convenienza all'eliminazione del cliente dal proprio portafoglio, portandolo all'abbandono.

Quando si parla di CRM, si può fare riferimento a¹⁵:

- **CRM come tecnologia**, ossia uno strumento software, di solito funzionante in Cloud, utilizzato per memorizzare tutti i dati fondamentali riferiti ai clienti e alle loro transazioni, che sarà poi impiegato per effettuare analisi

¹⁴ MARCHI L., MANCINI D. (a cura di), *Gestione informatica dei dati aziendali*, terza edizione, Franco Angeli, Milano, 2009, pag. 159.

¹⁵ Classificazione proposta in: <https://www.salesforce.com/it/learning-centre/crm/what-is-crm/>

e creare report utili all'alta direzione. (può essere anche definito "sistema CRM o CRM solution").

- **CRM come strategia**, ossia come filosofia di gestione basata sulla "centralità del cliente" (cliente-centrica) che fa riferimento alle modalità con le quali devono essere gestiti i rapporti con i clienti attuali e potenziali.
- **CRM come processo**, facendo riferimento ad un sistema che le aziende adottano per consolidare e gestire questi rapporti.

In questa sede, l'attenzione è concentrata solo sulla prima accezione, ossia quella riferita al Software utilizzato per la raccolta, rielaborazione ed analisi dei dati per singolo cliente in modo da ottenere una visione generale di questo, dei suoi comportamenti e poter fare previsioni sulle sue scelte future.

In generale, dati a cui si fa riferimento sono quelli anagrafici, di contatto (telefono, e-mail, indirizzo), dati circa le preferenze nella comunicazione, il numero dei contatti avuti con il cliente e le informazioni ottenute, le campagne di marketing che lo hanno interessato, gli acquisti effettuati.

Sulla base di ciò, sono state individuate 5 fasi nella gestione della relazione con i clienti:

- Identificazione e conoscenza del cliente
- Segmentazione dei clienti

- Contatti con i clienti, per capire quali sono le esigenze
- Personalizzazione dell'offerta con conseguente fidelizzazione
- Apprendimento *continuo* a partire dalle opinioni dei clienti.



Figura 2 Il CRM come sistema circolare.

(Fonte: De Luca A., CRM: Customer Relationship Management per la redditività dell'impresa).

Dunque, “il cuore del customer relationship management è il **customer database**, che contiene dati rilevanti per gestire con intelligenza le relazioni con i clienti. Queste relazioni si sviluppano in stadi successivi che disegnano un percorso circolare potenzialmente molto lungo: dal primo contatto alla fedeltà comportamentale e condivisa che si manifesta nel riacquisto sistematico”¹⁶.

¹⁶ MAURI C., *Il sistema informativo per il marketing. Dai dati alle risorse di conoscenza per la gestione delle relazioni con i clienti*, Giappichelli Editore, Torino, 2007, pag. 14.

In questi sistemi, tutti i processi che coinvolgono i clienti vengono integrati tra le funzioni di back office (contabilità, acquisti, produzione) e front office (marketing, vendite, assistenza) di modo che qualunque operatore delle differenti aree possa rispondere tempestivamente alle specifiche richieste del cliente. Di conseguenza il database dei clienti non contiene solo i dati anagrafici, ma anche le informazioni derivanti dall'area marketing, dall'area commerciale e customer care.

Dunque, nonostante tali sistemi si focalizzano sulle relazioni con i clienti, contengono una serie di dati che fanno da supporto ai diversi processi inter-organizzativi quali la contabilità, il marketing, la direzione, al fine di segmentare i clienti in base a criteri prestabiliti per poter personalizzare l'offerta¹⁷.

Si può notare come i dati potenzialmente da inserire possono essere molteplici e permetterebbero di ottenere una visione completa del cliente in modo da rendere così il rapporto personalizzabile, su misura del cliente stesso, delle sue caratteristiche e preferenze.

Per poter effettuare qualunque tipo di analisi sui clienti occorre avere a disposizione un database contenente i dati riferiti ai clienti che deve essere continuamente aggiornato e alimentato con le informazioni richieste. Solitamente i database contengono dati anagrafici e dati transazionali, riferiti al singolo cliente,

¹⁷ TAGLIAVINI M., RAVARINI A., SCIUTO D., *Sistemi per la gestione dell'informazione*, Apogeo, Milano, 2003, pag. 171.

e spesso si trovano sotto forma di dati elementari ma articolati in diversi campi che, uniti, formano un dato completo¹⁸.

Attraverso l'analisi di questa serie di dati si arriva solitamente a delineare un profilo di cliente (customer profiling), si identifica cioè un cliente "tipo" avente le caratteristiche più diffuse tra i clienti analizzati e sulla base di tali caratteristiche avviene la suddivisione in classi omogenee (clusterizzazione).

"Il customer profiling [...] è un concetto polivalente che comprende qualunque attività di esplorazione, descrizione, interpretazione e previsione di caratteri sociodemografici, percezioni, atteggiamenti, preferenze e comportamenti dei clienti. Il customer profiling ha due finalità principali:

- * Esplorativa – descrittiva
- * Interpretativa – predittiva.

Pur distinte, sono due finalità fortemente correlate¹⁹.

Nella prima fase si individuano, attraverso le tecniche di data mining, le associazioni tra i dati in base alla loro similitudine, all'omogeneità dei valori assunti dai dati e al tipo di informazione che possono offrire.

¹⁸ MAURI C., *Il sistema informativo per il marketing. Dai dati alle risorse di conoscenza per la gestione delle relazioni con i clienti*, Giappichelli Editore, Torino, 2007, pag. 112.

¹⁹ MAURI C., *Il sistema informativo per il marketing. Dai dati alle risorse di conoscenza per la gestione delle relazioni con i clienti*, Giappichelli Editore, Torino, 2007, pag. 113.

Nella fase predittiva, partendo dai gruppi di clienti individuati si va a calcolare la probabilità che certi eventi si verifichino o in che modo certe variabili possono mutare nel tempo. In base ai risultati di questa fase si formuleranno strategie adeguate in base agli obiettivi da perseguire e sarà anche possibile effettuare analisi degli scenari, per cercare di immaginare come alcune situazioni potranno evolversi e quali interventi poter effettuare per ogni diversa supposizione.

Parlando di sistemi informativi e CRM, un'altra importante accezione è quella di **Marketing Intelligence**. Si tratta di un sistema che, a differenza dei sistemi informativi tradizionali, accoglie solo i dati riferiti al marketing che non avranno un'utilità immediata ma che potrebbero essere utilizzati in futuro per prendere decisioni commerciali "più consapevoli" e "più intelligenti". I dati di cui si parla "possono essere dati relativi a trend socio – demografici, a clienti, concorrenti, a particolari eventi. Le modalità di raccolta dei dati sono varie: leggere libri, sfogliare pubblicazioni, navigare su internet, dialogare con i clienti, con i fornitori, con i distributori, con opinion leader"²⁰.

I vantaggi che il CRM offre alle aziende sono molteplici.

²⁰ MAURI C., *Il sistema informativo per il marketing. Dai dati alle risorse di conoscenza per la gestione delle relazioni con i clienti*, Giappichelli Editore, Torino, 2007, pag. 29.

Sicuramente uno dei punti di forza di un CRM è quello di avere una visione a 360 gradi del cliente in un unico luogo e, attraverso l'utilizzo di dashboard e schemi preimpostati si possono interrogare ed esplorare i dati, effettuare analisi statistiche e/o analizzare gli andamenti di determinate variabili.

Probabilmente il successo di questo tipo di software è dovuto all'introduzione, negli ultimi anni, di CRM disponibili in piattaforme Cloud, utilizzabili senza dover installare il programma in ogni singolo computer e permettendo la fruibilità di informazioni aggiornate in tempo reale da parte di tutti gli utenti interessati, quindi maggiore collaborazione tra le diverse aree aziendali (che verrebbero così a comunicare tra loro) e rendendo più snelli i processi.

Inoltre, un importante vantaggio è dato dal fatto che questi sistemi sono più immediati perché non richiedono alcuna installazione e vengono aggiornati automaticamente alle ultime versioni disponibili ed a costi contenuti.

Infine, offrono la possibilità di accesso anche oltre i "confini" aziendali quindi in qualsiasi luogo e in ogni momento, su ogni tipo di dispositivo dotato di accesso ad Internet.

I vantaggi del CRM possono anche essere ricondotti alle attività più strettamente operative. Tra questi troviamo il fatto che questi sistemi permettono di avere una visione "evolutiva" del cliente poiché si tiene traccia dei dati che man mano vengono aggiunti nel sistema, delle modifiche effettuate e indicano anche il

soggetto che ha effettuato le modifiche, il giorno e l'ora in cui la modifica viene effettuata.

Altro vantaggio di tipo operativo è dato dal fatto che i dati sui clienti vengono integrati tra le diverse funzioni aziendali quindi si ha la disponibilità in un unico "luogo" di tutte le informazioni riferite ad un cliente specifico, siano esse di provenienza, di redditività, di marketing o di customer care.

CAPITOLO 2: IL DATA WAREHOUSE

PREMESSA

Considerando l'importanza attribuita alla conoscenza e al ruolo sempre più strategico che il cliente riveste per l'azienda, si punta sempre più a studiarne i comportamenti e le preferenze attraverso una serie di variabili. I dati raccolti nei sistemi informativi saranno quindi selezionati e formeranno un data warehouse cioè un magazzino in cui si trovano tutti i dati riferiti ad un certo tema, in questo caso sul cliente, che permetteranno di svolgere l'analisi.

Attraverso gli OLAP si potranno effettuare navigazioni sui dati aggregati e tramite strumenti di ETL sarà possibile estrarre i dati, individuare errori, ripulirli e rielaborarli per poi inserirli nei software di analisi e individuare quali informazioni saranno strategicamente utili al processo decisionale.

2.1 COS'È UN DATAWAREHOUSE

Il data warehouse è l'elemento centrale di un sistema di Business intelligence²¹.

Può essere definito come “la base di dati informazionali, che raccoglie in un unico magazzino (warehouse significa magazzino) tutti i dati di interesse per l'azienda e

²¹ “Per **Business Intelligence** si intendono tutti i processi e gli strumenti attraverso i quali un'azienda riesce a raccogliere dati di diversa natura per analizzarli e trarne decisioni strategiche.”
<https://www.extrasys.it/it/smartblog/che-cosa-%C3%A8-la-business-intelligence>

relativi alla sua attività di business, sintetizzandoli, integrandoli e strutturandoli in modo tale da rendere agevole ed efficace la ricerca di informazioni.” (Pighin e Marzona, 2011)

Obiettivo del data warehouse è quindi quello di raccogliere dati potenzialmente interessanti, relativi ad un processo/un’attività che si vuole prendere in esame per ottenere informazioni utili al processo decisionale dei manager e/o degli analisti.

I dati contenuti nel datawarehouse possono essere attuali oppure riferiti a periodi precedenti e possono provenire dai sistemi informativi delle diverse aree aziendali quali Acquisti, Logistica, Produzione, Vendite... ma anche da altre fonti (si veda il paragrafo 1.2).

C’è differenza tra il data warehouse e il data warehousing; quest’ultimo infatti può essere inteso secondo due accezioni: da un lato, è l’insieme di metodologie del tipo ETL (Extract, Transform and Load) che estraggono i dati dalle fonti, li trasformano, controllano, puliscono e riorganizzano per consolidare dati informativi consistenti, e dall’altro come l’insieme di strumenti di analisi che supportano i processi decisionali valutando i dati informativi sulla base delle richieste formulate dall’utente²².

²² Cfr. PIGHIN M., MARZONA A., *Sistemi informativi aziendali, ERP e sistemi di data analysis*, Terza edizione, Pearson, 2018, pag. 265.

È opportuno anche sottolineare la differenza tra il concetto di **Database** che rappresenta uno strumento che raccoglie tutti i dati provenienti dai sistemi informativi aziendali e li registra in tempo reale in un'unica piattaforma, e il **datawarehouse** ossia un database "speciale" in cui sono presenti solo le informazioni necessarie per effettuare l'analisi, che funzionano in base a sistemi di tipo OLAP²³ (On Line Analytical Processing). Altra differenza è legata al fatto che non solo il DW è più diretto all'analisi perché contiene dati selezionati, ma anche al fatto che il database viene aggiornato in tempo reale perciò si potrebbero avere problemi nel momento in cui i dati vengono aggiornati: in questo senso sarebbe meglio optare per la creazione di un DW perché permette di avere un'istantanea sui dati riferiti ad un certo periodo di tempo²⁴.

2.2 ELEMENTI COSTITUTIVI DI UN DATA WAREHOUSE

All'interno di un data warehouse si trovano 3 tipologie di dati:

- 1- Dati interni: sono quei dati presenti nei sistemi operazionali e raccolti tramite software gestionali del tipo ERP (Enterprise Resource Planning) e che riguardano clienti, fornitori, vendite, acquisti, risorse umane.

²³ Sul concetto di OLAP si rimanda al paragrafo 1.7

²⁴ Rif. <https://www.bucap.it/news/approfondimenti-tematici/gestione-del-magazzino/database-data-warehouse-principali-differenze.htm>

- 2- Dati esterni: si tratta principalmente di banche dati messe a disposizione da società di servizi specializzate che permettono di integrare le informazioni disponibili internamente. Si tratta di dati riguardanti vendite dei competitor, quote di mercato, indicatori economici e finanziari riferiti all'andamento del settore...
- 3- Dati personali: si tratta di "informazioni e valutazioni personali" elaborate dai decision maker, spesso frutto di rielaborazioni ed integrazioni tra dati esterni e interni.

Altro elemento costitutivo di un data warehouse sono le tabelle delle dimensioni e le tabelle dei fatti.

Le **tabelle delle dimensioni** contengono le informazioni relative alle entità primarie presenti nel data warehouse. Con entità primarie s'intendono gli elementi fondanti dell'attività aziendale: unità di business, punti vendita, clienti, prodotti, tempo e così via. All'interno di ogni tabella di dimensione possiamo trovare strutture di tipo **gerarchico**²⁵, che sono fondamentali nel momento in cui si andranno a sviluppare le analisi OLAP.

Le **tabelle dei fatti** invece, contengono le caratteristiche e le misure riferite alle dimensioni oggetto di analisi e si collegano alle tabelle delle dimensioni primarie.

²⁵ Un esempio di gerarchia è rappresentato dalla dimensione "data" che si articola in giorni, settimane, mesi, anni.

Una tabella dei fatti può essere legata a n tabelle delle dimensioni e questo si rappresenta attraverso i “cubi di dati multidimensionali” (c. d. Cubi OLAP). I cubi multidimensionali costituiscono una naturale estensione dei fogli elettronici, e questi ultimi possono essere infatti interpretati come cubi bidimensionali²⁶.

2.3 I DATA MART

I dati raccolti all'interno del sistema informativo aziendale e che popolano il data warehouse possono essere riorganizzati e/o riaggregati, sulla base degli obiettivi aziendali da perseguire, attraverso l'utilizzo dei data mart.

I **data mart** rappresentano sottoinsiemi di data warehouse che accolgono i dati di una sola attività aziendale, di un settore, di un singolo processo o di un tema, particolarmente utili per agevolare le analisi, e quindi le decisioni, in quanto trattano una minore quantità di dati ma soprattutto dati mirati all'analisi stessa. Poiché i dati contenuti in questi sistemi sono decisamente minori rispetto al magazzino dati principale, si dice che i data mart permettono di eseguire query complesse in tempi più brevi e garantiscono una migliore qualità delle analisi dovuta al fatto che i dati saranno più facilmente controllabili e quindi anche più precisi.

²⁶ VERCELLIS C., *Business Intelligence, modelli matematici e sistemi per le decisioni*, McGraw-Hill, 2006, pag. 57.

“Grazie ai data mart, gli analisti e i manager aziendali possono definire le angolature dalle quali osservare l’insieme (spesso incoerente) di informazioni rappresentato dal data warehouse²⁷”. Infatti i dati contenuti nel “magazzino” non sempre sono tutti utili ai fini dell’analisi e perciò dovranno essere selezionati per area tematica in modo da avere un’idea più approfondita di alcuni aspetti determinanti che meritano particolare attenzione.

Inoltre, i data mart in quanto parti di datawarehouse più ampi, rappresentano “l’unità informativa di base per l’analisi con i sistemi di BI”²⁸.

A partire da un data warehouse quindi è possibile estrarre moltissimi data mart, tanti quanti sono gli obiettivi aziendali e quindi le analisi necessarie a perseguirli, e proprio come i data warehouse è possibile crearli anche a partire da fonti di dati esterne.

Anche nella creazione dei data mart si può seguire infatti una logica top down in cui i data mart costituiscono vere e proprie “scomposizioni” di data warehouse, oppure è possibile mettere insieme una serie di dati interni ed esterni a formare i data mart tematici che, uniti, formano il data warehouse (approccio bottom up).

²⁷ <https://www.bucap.it/news/approfondimenti-tematici/gestione-del-magazzino/definizione-data-mart.htm>

²⁸ <https://www.bucap.it/news/approfondimenti-tematici/gestione-del-magazzino/definizione-data-mart.htm>

Una differenza rispetto al data warehouse sta nel fatto che mentre quest'ultimo contiene i dati dell'intera azienda, i data mart sono in grado di rispondere alle esigenze di una singola divisione per cui possono essere utilizzati anche come strumento di una politica di gestione delle informazioni che prevede l'accessibilità dei dati solo ad alcune categorie di utenti, controllando chi può accedervi e ciò accade principalmente con i dati sensibili o comunque nelle grandi aziende in cui vi è una mole di dati non indifferente e ciò potrebbe creare confusione.

In conclusione, i vantaggi offerti dal data mart sono principalmente da ricercare nelle differenze rispetto al data warehouse e sono sintetizzabili come segue:

- Volumi di dati ridotti, riguardanti uno specifico tema
- facile accesso ai dati richiesti di frequente
- più semplici da implementare rispetto al Datawarehouse e meno costosi
- maggiore flessibilità: una modifica del modello può avvenire rapidamente perché le dimensioni sono ridotte
- i dati sono fruibili solo ai diretti interessati poiché sono già suddivisi per aree tematiche
- è specifico per le esigenze di alcune categorie di utenti quindi può velocizzare i processi decisionali

- Contiene i dati riferiti agli eventi storici che permetterebbero all'analista di determinare i trend di un certo dato²⁹.

Vercellis afferma che di solito le aziende preferiscono “progettare e realizzare in modo incrementale una serie di **data mart** tra loro integrati piuttosto che un data warehouse centrale in modo da ridurre i tempi e le incertezze legate al progetto”³⁰.

A mio avviso, dato che ogni azienda ha degli asset differenti e delle caratteristiche *firm specific*, avrà esigenze altrettanto diverse sia in termini organizzativi che di conoscenza, motivo per il quale non è detto che parte necessariamente da tanti data mart integrati ma si potrebbe preferire la creazione di un data warehouse centrale che sarà poi scomposto secondo i diversi temi. (Si vedano i modelli di Inmon e Kimball nel successivo paragrafo)

²⁹ Cfr. <https://www.guru99.com/data-mart-tutorial.html>

³⁰ VERCELLIS C., *Business Intelligence, modelli matematici e sistemi per le decisioni*, McGraw-Hill, 2006, pag. 51.

2.4 ARCHITETTURE DI DATA WAREHOUSE

I sistemi di data warehousing sono costituiti dalle basi di dati, diverse per finalità, struttura e tipologia di dati contenuti, che si dispongono su diversi livelli come illustrato in figura.

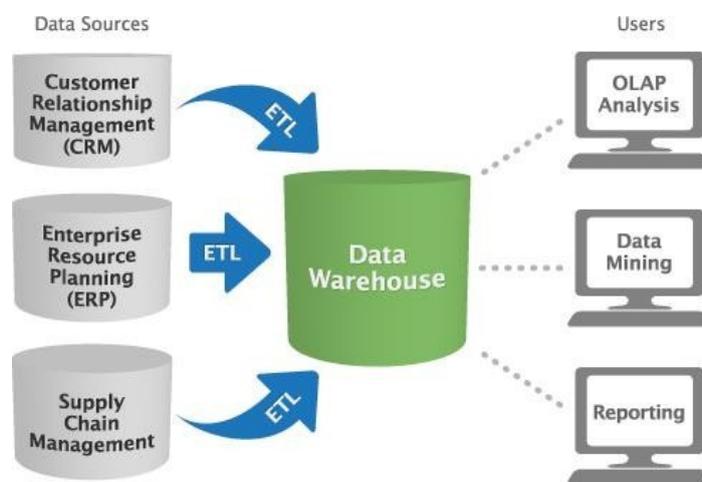


Figura 3. Il data warehouse.

Si nota che il data warehouse costituisce l'elemento centrale mentre a monte sono presenti quegli elementi che permettono la trasformazione dei dati dalle sorgenti al modello multidimensionale progettato. A valle invece ci sono gli eventuali data mart, articolati su gerarchie più o meno complesse e realizzati con lo scopo principale di rendere più efficienti i processi di analisi dei dati³¹.

³¹ Si veda: PIGHIN M., MARZONA A., *Sistemi informativi aziendali, ERP e sistemi di data analysis*, Terza edizione, Pearson, 2018, pag. 286.

Il modello di data warehouse sarà differente a seconda degli obiettivi da raggiungere, della complessità di analisi e della quantità di dati da trattare.

Solitamente si articola il DW su 3 livelli:

- 1) Il primo livello è costituito da strumenti di *back-end*, ossia dalle fonti di dati che permettono il popolamento del datawarehouse e dai relativi strumenti di ETL (Estrazione, Trasformazione, Caricamento) poiché i dati provengono da diverse applicazioni;
- 2) Il livello costituito dal dw e dagli eventuali data mart tematici articolati in modo più o meno gerarchico;
- 3) Nel terzo livello si trovano gli strumenti di analisi interattiva che permettono la navigazione sugli ipercubi e la rappresentazione chiara e intuitiva dei dati ai diversi livelli di aggregazione e confronto³². A questo livello troviamo quindi “le applicazioni di business intelligence e di supporto alle decisioni che costituiscono i *front end* e permettono ai knowledge worker di svolgere le analisi e visualizzare i risultati”³³.

³² PIGHIN M., MARZONA A., *Sistemi informativi aziendali, ERP e sistemi di data analysis*, Terza edizione, Pearson, 2018, pag. 287.

³³ VERCELLIS C., *Business Intelligence, modelli matematici e sistemi per le decisioni*, McGraw-Hill, 2006, pag. 52.

Il livello centrale costituisce “il cuore del sottosistema informativo direzionale”³⁴ contenente i dati organizzati secondo 3 modalità: attuali di dettaglio, aggregati e storici di dettaglio.

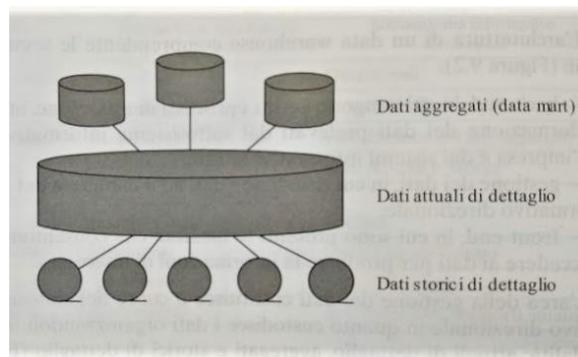


Figura 4. L'area gestione dati del datawarehouse (Fonte: Micromarketing, pag.200).

Di solito i dati attuali sono quelli più voluminosi perciò bisogna effettuare un processo di selezione dei dati tenendo conto non solo delle finalità da soddisfare (qui il riferimento è all'efficacia), ma anche della quantità di dati da considerare (efficienza).

³⁴ LUGLI G., ZILIANI C., *Micromarketing, creare valore con le informazioni di cliente*, UTET, 2004, pag. 200.

“I dati contenuti in questa area sono già stati sottoposti ad operazioni di filtraggio, integrazione tra fonti diverse e trasformazione”³⁵ prima ancora di popolare il data warehouse, ma restano comunque tanti perché più importanti ai fini analitici.

I dati aggregati sono quelli messi a disposizione dell'utente per velocizzare la ricerca di informazioni (dato che non vengono né analizzati né rielaborati), anche se in questo modo i dati perdono il loro livello di dettaglio. In genere i dati vengono aggregati con tecniche di data mart ossia per aree tematiche³⁶.

I dati storici di dettaglio sono invece quelli che occupano meno spazio in quanto vengono archiviati ed utilizzati solo in fase di confronto con la situazione attuale.

La realizzazione di un data warehouse può seguire diverse logiche³⁷:

- a) Top down (modello di Inmon): si forma il data warehouse completo e successivamente si scompone in diversi data mart. Sicuramente sarà più preciso e dettagliato ma comporta il sostenimento di elevati costi e richiede più tempo.
- b) Bottom up (modello di Kimball): si parte dal particolare, ossia si costituisce prima il data mart che verrà poi ampliato fino a formare il datawarehouse.

³⁵ LUGLI G., ZILIANI C., *Micromarketing, creare valore con le informazioni di cliente*, UTET, 2004, pag. 200.

³⁶ LUGLI G., ZILIANI C., *Micromarketing, creare valore con le informazioni di cliente*, UTET, 2004, pag. 199.

³⁷ Si veda: REZZANI A., *Business Intelligence, processi, metodi, utilizzo in azienda*, Apogeo education, Maggioli editore, 2012, pag. 99 e VERCELLIS C., *Business Intelligence, modelli matematici e sistemi per le decisioni*, McGraw-Hill, 2006, pag. 53-54.

Da un lato, comporta brevi tempi di implementazione e costi più contenuti rispetto al modello precedente, ma dall'altro potrebbe portare a deviare l'attenzione dall'obiettivo principale di analisi.

- c) Mista: consiste nello sviluppare un progetto complessivo di dw che permette una visione d'insieme, e di procedere quindi alla realizzazione di prototipi successivi corrispondenti a parti diverse del sistema.

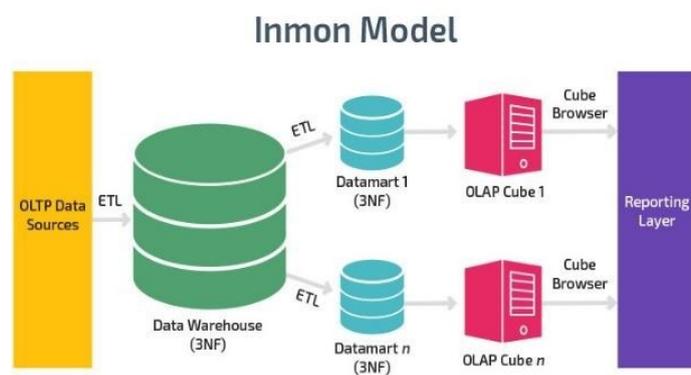


Figura 5 Il modello di Data Warehouse secondo Inmon. (Fonte: <https://panoply.io/data-warehouse-guide/data-mart-vs-data-warehouse/>)

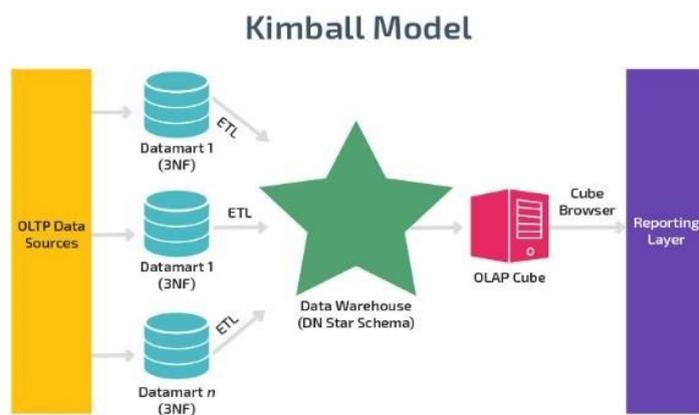


Figura 6 Il modello di Data Warehouse secondo Kimball (Fonte: <https://panoply.io/data-warehouse-guide/data-mart-vs-data-warehouse/>)

In ogni caso, qualunque sia la scelta architetture, gli step da seguire per realizzare un data warehouse, o un data mart, possono essere schematizzati come segue³⁸:

- Identificazione del processo organizzativo da rappresentare nel datawarehouse, (processo di vendita, la logistica, il controllo di gestione...);
- Definizione della granularità, ossia del livello di dettaglio dei dati da analizzare e il loro livello di aggregazione;
- Scelta delle misure rilevanti da inserire nelle tabelle dei fatti;
- Determinazione delle dimensioni per ciascuna tabella dei fatti.

Poiché alla base di un dw i dati sono classificati secondo una gerarchia, le tabelle dei dati e delle dimensioni potranno essere rappresentate tramite il **Dimensional Fact Model** ossia uno schema complesso formato da tabelle di fatti a cui vengono collegate, tramite frecce, le dimensioni di base e gli attributi (anche derivati) della singola dimensione.

³⁸ VERCELLIS C., *Business Intelligence, modelli matematici e sistemi per le decisioni*, McGraw-Hill, 2006, pag. 54.

Oltre al modello DFM, i dati possono essere rappresentati tramite schemi che sottolineano l'importanza della relazione e, a seconda della complessità delle relazioni stesse, possono essere di 3 tipologie: a stella, a fiocco di neve o a costellazione.

SCHEMA A STELLA

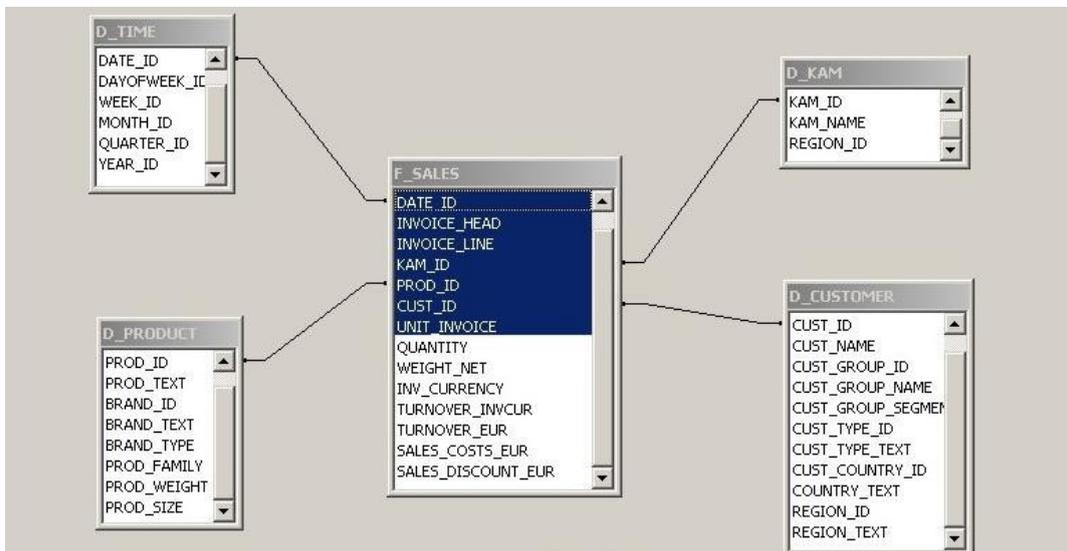


Figura 7 Schema a stella (<https://etl-tools.info/en/bi/datawarehouse/star-schema.htm>)

In questo tipo di schema si pone al centro la tabella del fatto e le tabelle delle dimensioni intorno, legate alla tabella centrale da un'unica freccia indicante la relazione. La caratteristica principale è che in questo tipo di schema le tabelle sono denormalizzate: sono presenti ridondanze e gli elementi gerarchici della

dimensione sono memorizzati nella stessa tabella per cui non si riescono ad individuare le gerarchie profonde.³⁹

SCHEMA A FIOCCO DI NEVE

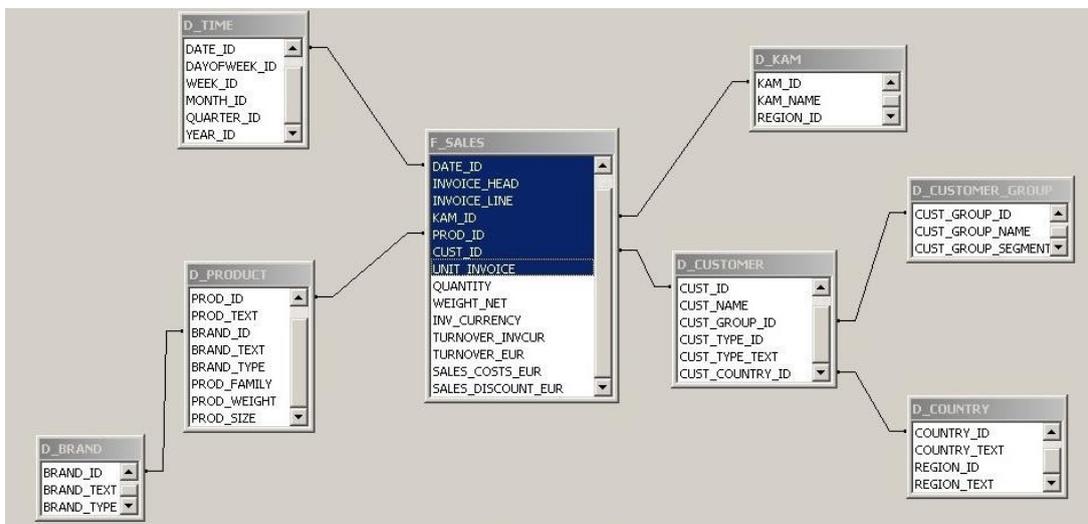


Figura 8 Schema a fiocco di neve (https://etl-tools.info/en/bi/datawarehouse_snowflake-schema.htm)

Anche in questo caso vi è la tabella del fatto al centro e le relative tabelle delle dimensioni collegate ma a queste ultime vengono collegate ulteriori tabelle che rappresentano le dimensioni secondarie. Un'altra differenza si riscontra nel fatto che le tabelle qui sono normalizzate perciò sono già predisposte per i database relazionali.

³⁹ PIGHIN M., MARZONA A., *Sistemi informativi aziendali, ERP e sistemi di data analysis*, Terza edizione, Pearson, 2018, pag. 294.

SCHEMA A COSTELLAZIONE

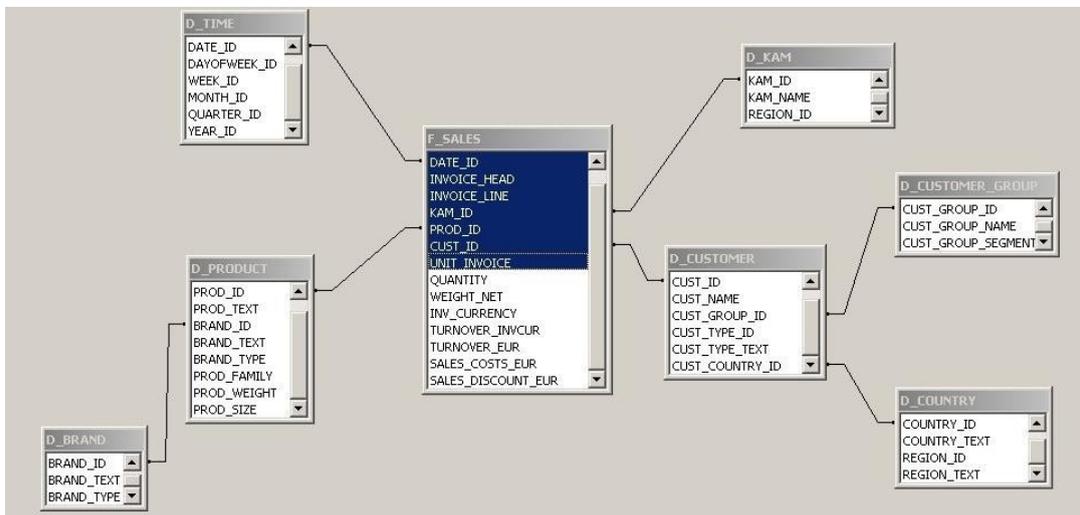


Figura 9 Schema a costellazione (https://etl-tools.info/en/bi/datawarehouse_constellation-schema.htm)

Si costruisce uno schema a costellazione quando le tabelle dei fatti condividono le stesse dimensioni (è il caso della tabella dimensionale “Data” che può essere condivisa tra più tabelle). Ovviamente questo è uno schema più complesso da gestire.

In conclusione, a mio avviso è preferibile questa seconda tipologia di schemi perché risultano più facilmente comprensibili e soprattutto hanno un maggiore impatto diretto rispetto alle tabelle del Dimensional Fact Model.

2.5 QUALITÀ DEI DATI

“La qualità dei dati è un requisito fondamentale per l'intero sistema informativo”⁴⁰. Dopo aver aggregato i dati nel database infatti, l'analista si deve preoccupare di effettuare una serie di azioni volte alla “selezione” dei dati, alla loro qualità.

Nel momento in cui si vanno a trasferire dai sistemi operazionali al datawarehouse, i dati devono essere esaminati e corretti per ridurre al minimo gli errori, la cui presenza potrebbe inficiare sulle analisi e quindi sui risultati. Per questo motivo intervenire sulla qualità dei dati “aumenta evidentemente il valore e la significatività delle analisi di Business Intelligence sviluppate a partire da un data warehouse” (Vercellis 2006).

In particolare, gli analisti di dati e coloro che si occupano di progettazione e aggiornamento del data Warehouse, devono preoccuparsi della qualità poiché i dati potrebbero essere non corretti oppure non aggiornati o ancora, inesistenti (vedi tabella).

⁴⁰ REZZANI A., *Business Intelligence, processi, metodi, utilizzo in azienda*, Apogeo education, Maggioli editore, 2012, pag. 26.

INCONVENIENTE	CAUSA	RIMEDIO
Dati non corretti	Dati raccolti senza adeguata cura Dati inseriti in modo errato Dati modificati in modo incontrollato	Controllo sistematico dei dati in ingresso Automazione di data entry Predisposizione di un programma di sicurezza per accessi e modifiche
Dati non aggiornati	La raccolta dei dati non è allineata alle esigenze degli utilizzatori	Meccanismi tempestivi di raccolta e aggiornamento Utilizzo del web per ottenere dati aggiornati
Dati inesistenti	Mancata raccolta dei dati richiesti	Identificazione dei dati necessari mediante accurate analisi preliminari e stima dei dati mancanti

Tabella 1. Cause e possibili rimedi alla scarsa qualità dei dati (Fonte: Vercellis 2006, pag. 51).

Le motivazioni circa l'importanza della qualità sono da ricercarsi nel fatto che eventuali errori presenti all'interno del database si ripercuotono poi sull'analisi vera e propria, quindi sui risultati (che verrebbero così ad essere inficiati) e sulle conseguenti decisioni che potrebbero portare non pochi problemi in termini di peggioramento della performance.

Gli errori possono essere riconducibili a:

- Errori manuali

- Errori nel software di gestione dei dati
- Errori di progettazione delle basi dati
- Errori nei processi di gestione dei dati⁴¹.

Per rimediare a questi errori è necessario che i dati abbiano determinate caratteristiche:

- * Accuratezza: i dati devono essere il più possibile corretti ed essere conformi ai valori reali ossia devono essere compresi in un intervallo ben definito;
- * Completezza: è legata al punto precedente, infatti per non compromettere l'accuratezza delle analisi non devono esserci molti valori mancanti, soprattutto se questi sono considerati necessari.
- * Consistenza: non devono esserci contraddizioni nella forma e nel contenuto dei dati che provengono da uno stesso o da diversi database integrati;
- * Attualità: sarebbe auspicabile analizzare dati il più possibile aggiornati, anche in base agli obiettivi delle analisi che si vogliono effettuare (nulla toglie di poter utilizzare dati storici);

⁴¹ REZZANI A., *Business Intelligence, processi, metodi, utilizzo in azienda*, Apogeo education, Maggioli editore, 2012, pag. 26.

- * Non ridondanza⁴²: bisogna cercare di evitare repliche e ridondanze per una questione di spazio (memoria) ma principalmente per non interferire nei risultati;
- * Rilevanza: i dati devono essere significativi ai fini dell'analisi e devono dare un valore aggiunto all'analisi stessa;
- * Interpretabilità: gli analisti devono essere in grado di comprendere l'importanza dei dati e il loro significato;
- * Accessibilità: i dati devono essere a completa disposizione degli analisti e dei manager, cioè questi soggetti devono poter usufruire dei dati in maniera semplice e veloce ed in qualsiasi momento.

Altri problemi legati alla scarsa qualità dei dati sono costituiti dal sostenimento di maggiori costi e dispendio di tempo per la correzione degli errori, la perdita di opportunità che potrebbero non essere più colte in seguito, la riallocazione delle risorse finanziarie e la successiva modifica delle strategie messe in atto dai manager⁴³.

⁴² La **ridondanza** fa riferimento alla presenza di dati ripetuti. Ciò si verifica principalmente quando l'azienda si organizza per funzioni/divisioni che raccolgono gli stessi tipi di dati all'interno di file distinti. La ridondanza porta ad uno spreco di risorse ma soprattutto all'incoerenza dei dati nel momento in cui lo stesso attributo assume diversi valori. (Management dei sistemi informativi p.81)

⁴³ Si veda: REZZANI A., *Business Intelligence, Processi metodi e utilizzo in azienda*, Apogeo, 2012, pag. 26.

Ovviamente, se un database è stato progettato in modo idoneo e si sono stabiliti gli standard a livello aziendale, i dati saranno più precisi e coerenti perciò di migliore qualità.

2.6 GLI STRUMENTI DI ETL

ETL è l'acronimo di Extract, Transform and Load. Consiste quindi nell'insieme delle procedure di estrazione, trasformazione e caricamento dei dati nel datawarehouse.

Probabilmente l'ETL è uno degli elementi più complessi di un sistema di Business Intelligence perché con esso "si alimenta il data warehouse e si compiono le operazioni di pulizia, integrazione e uniformazione dei dati"⁴⁴.

Nella fase di estrazione vengono estrapolati i dati di interesse (prima estrazione) dalle diverse fonti disponibili, interne ed esterne, andando a costituire il magazzino dati. Le estrazioni successive serviranno ad aggiornare il DW in base ai dati che si renderanno disponibili con il trascorrere del tempo.

La trasformazione fa riferimento all'attività di pulitura dei dati e la correzione di eventuali errori. A volte si sostituiscono i termini errati oppure si modificano termini simili, si aggregano dati omogenei oppure si rendono omogenei i dati che

⁴⁴ REZZANI A., *Business Intelligence, Processi metodi e utilizzo in azienda*, Apogeo, 2012, pag. 199.

provengono da fonti diverse. Si potrebbero inserire anche delle regole automatiche per la correzione degli errori⁴⁵.

L'ultima è la fase di caricamento che consiste nell'inserimento dei dati all'interno dei DW per metterli a disposizione degli analisti e dei sistemi di supporto alle decisioni.

Dopo aver effettuato il caricamento dei dati e identificato i collegamenti e le gerarchie tra di essi, il punto di partenza delle analisi è costituito dalle operazioni OLAP (On Line Analytical Processing).

2.7 OLAP

Come già ribadito, il data warehouse è costituito da un'enorme mole di dati che potranno essere analizzati attraverso gli strumenti di Business Intelligence utili a supportare le decisioni dei manager. In questo senso, uno strumento particolarmente utile all'analista risulta essere l'operatore OLAP che permette la navigazione attraverso i dati e facilita il reperimento delle informazioni di cui si necessita in tempi abbastanza brevi. La struttura del data warehouse con i suoi elementi costitutivi rappresenta il punto di partenza per le analisi OLAP.

⁴⁵ VERCELLIS C., *Business Intelligence, modelli matematici e sistemi per le decisioni*, Mc-Graw Hill, 2006, p. 54.

Il DW è infatti basato su un modello dati multidimensionale: i dati sono contenuti in tabelle caratterizzate da moltissime righe e colonne e sono collegati tra loro, anche tramite gerarchie. Non solo. I dati contenuti nelle diverse tabelle dei fatti e delle dimensioni possono essere legati anche ad altre tabelle: tutto questo “insieme” di tabelle può essere rappresentato attraverso i cubi di dati multidimensionali⁴⁶ che legano una tabella di fatti a n tabelle delle dimensioni.

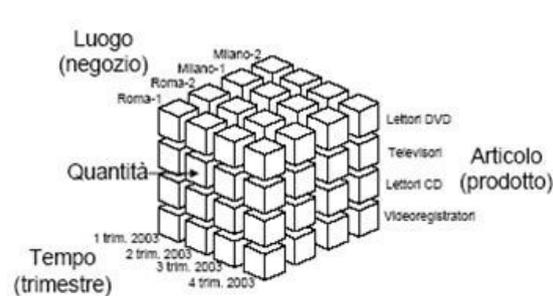


Figura 10. esempio di cubo multidimensionale dei dati.

I dati sono perciò visti sotto forma di ipercubi in cui ad ogni dimensione (altezza, larghezza, profondità) possono corrispondere una serie di tabelle contenenti le caratteristiche dei dati.

L'ipercubo deve contenere:

- il **fatto**, ossia l'elemento oggetto di indagine;
- le **dimensioni**, cioè gli attributi che si sceglie di analizzare su quel fatto;

⁴⁶ I cubi multidimensionali costituiscono una naturale estensione dei fogli elettronici, che possiamo infatti interpretare come cubi bidimensionali. (Vercellis p. 57)

- le **misure** che danno valore agli eventi misurati.

Una volta organizzati secondo questa struttura, i dati saranno disponibili alle analisi, permetteranno di effettuare ricerche avanzate e saranno aggregabili o disaggregabili, a seconda delle esigenze.

Un OLAP può essere perciò definito come “l’insieme dei sottosistemi informativi aziendali pensati per l’analisi interattiva dei dati, ottimizzati per garantire la massima efficienza nell’elaborazione dei dati di sintesi e la massima flessibilità nelle interrogazioni”⁴⁷.

In sostanza, si tratta di strumenti interattivi che permettono la navigazione sui dati e la verifica delle ipotesi formulate dall’utente.

Possono basarsi su sistemi di sola lettura oppure sistemi che agevolano l’aggregazione dei dati, con strutture orientate agli oggetti di analisi.

I principali operatori OLAP sono:

- **drill down**: “scendere in profondità” ossia arrivare ad un maggior livello di dettaglio delle informazioni;
- **roll up**: è l’inverso del precedente, ossia è la tecnica con la quale si aggregano i dati per averne una dimensione più generale;

⁴⁷ PIGHIN M., MARZONA A., *Sistemi informativi aziendali, ERP e sistemi di data analysis*, Terza edizione, Pearson, 2018, pag. 31.

- **slice**: operazione che consente di ridurre la dimensionalità dei dati ad una sola porzione di essi, fissata una dimensione;
- **dice**: fissate più dimensioni, si limita l'insieme dei dati applicando un filtro;
- **pivot**: ruota l'ipercubo di analisi scambiando le dimensioni per ottenere una diversa vista sui dati.

L'analisi OLAP viene considerata come la principale modalità di fruizione delle informazioni presenti sui datawarehouse perché consente di navigare sui dati e permette anche di vedere gli stessi dati in modi diversi, proprio attraverso l'uso delle dimensioni multiple. Inoltre, permette agli utenti di ottenere risposte rapide a domande specifiche in un tempo piuttosto breve anche quando i dati sono memorizzati in database di grandi dimensioni⁴⁸.

⁴⁸ Laudon K. C., Laudon J. P., Morabito V., Pennarola F., Management dei sistemi informativi vol. 1 Fondamenti, Terza edizione, Pearson Prentice Hall, 2010.

CAPITOLO 3: IL DATA MINING

3.1 IN COSA CONSISTE IL DATA MINING

Finora si è parlato del datawarehouse e della sua importante funzione di raccolta di una grande quantità di dati che negli anni passati era impensabile.

Questa grande mole di dati utile ad effettuare analisi di Business Intelligence potrebbe però non essere sufficiente a descrivere certi fenomeni o comunque, se non ben utilizzata, potrebbe restare come una semplice base dati che contiene informazioni ma che non genera alcuna conoscenza. L'obiettivo ultimo di qualsiasi analisi è infatti proprio quello di generare conoscenza utile ai manager per poter prendere le decisioni migliori.

I sistemi di data warehouse hanno fornito una prima risposta all'esigenza informativa delle aziende: hanno permesso di riorganizzare i dati su cui è possibile condurre analisi in modo più semplice, veloce, controllato, facilitando il processo di decisione⁴⁹. Ma il data warehouse e le collegate analisi OLAP permettono di effettuare analisi dei dati partendo dalle ipotesi formulate dagli utenti, e questo significa che le analisi saranno vincolate da quelle che sono le conoscenze e le

⁴⁹ PIGHIN M., MARZONA A., *Sistemi informativi aziendali, ERP e sistemi di data analysis*, Terza edizione, Pearson, 2018, pag. 320.

esperienze dell'utente stesso. Infatti, come alcuni autori ammettono (Laudon et al. 2010), il data mining è più orientato alla scoperta rispetto alle analisi OLAP.

In questo senso gli strumenti di data mining si sono rivelati necessari, o comunque sempre più diffusi, per poter condurre analisi più approfondite sui dati, aumentare il grado di obiettività e precisione delle analisi stesse e considerare anche aspetti che altrimenti verrebbero sottovalutati o addirittura non valutati.

Il **data mining** può essere definito come un insieme di tecniche, modelli e strumenti di tipo statistico-matematico, utili ad individuare *automaticamente* le relazioni nascoste tra le variabili prese in esame e ad incrementare la capacità di previsione dell'azienda. In particolare, attraverso il data mining si analizzano campioni di dati riferiti al passato allo scopo di individuare le regole e gli algoritmi che consentirebbero di trarre conclusioni generali su tutta la popolazione⁵⁰ e che potranno essere utilizzati per poter costruire modelli previsionali.

Bisogna fare una precisazione circa la fase predittiva: è ovvio che "il data mining non è in grado di fare previsioni su nulla; semplicemente offre un'indicazione probabilistica del verificarsi di un evento"⁵¹.

⁵⁰ Cfr. VERCELLIS C., *Business Intelligence, modelli matematici e sistemi per le decisioni*, Mc-Graw Hill, 2006, p. 78.

⁵¹ REZZANI A., *Business Intelligence, Processi metodi e utilizzo in azienda*, Apogeo, 2012, pag. 47.

3.2 CREAZIONE DI UN SISTEMA DI DATA MINING

Il data mining può anche essere definito come un processo volto ad ottenere “una conoscenza sfruttabile per vari fini”: l'intero processo di conoscenza viene chiamato Knowledge Discovery in Databases (KDD) che comprende anche altre attività oltre a quella di data mining in senso stretto⁵².

Le fasi di cui si compone un processo di KDD sono:

- **identificazione dell'obiettivo** che si vuole raggiungere (è la fase in cui porre maggiore attenzione dato che una errata formulazione degli obiettivi potrebbe portare all'inutilità dell'analisi);
- **preselezione** dei dati utili a raggiungere l'obiettivo;
- **pulizia dei dati e preelaborazione**: dopo aver cercato di aumentare la qualità del data warehouse bisogna effettuare un'ulteriore selezione fra dati validi e inutili, ma anche la scelta di come trattare i campi incompleti o vuoti e la selezione definitiva delle informazioni fondamentali per il modello di riferimento;
- **trasformazione**: il formato con il quale sono rappresentati i dati deve essere idoneo all'attività di mining;

⁵² <http://www.intelligenzaartificiale.it/data-mining/>

- **data mining** (in senso stretto): costituisce l'analisi vera e propria. Si sceglie il software migliore a seconda degli obiettivi e delle analisi da effettuare e si individua l'algoritmo da utilizzare. Di solito questa fase si compone di più passaggi, che vengono più volte ripetuti per affinare la procedura e verificare man mano i risultati raggiunti;
- **interpretazione dei risultati**: le funzioni di mining riportano ciò che è considerato notevole dal punto di vista degli algoritmi utilizzati; questo però non significa che tutto ciò che è importante per l'algoritmo è considerato tale ai fini dell'analisi che si sta svolgendo. In questa fase bisogna selezionare solo i dati utili all'analisi e si andrà a valutare se l'obiettivo è raggiunto. Se così non fosse, si procede con la revisione (ed eventuale modifica) delle fasi precedenti;
- **presentazione dei risultati**: la conoscenza che deriva da queste fasi dovrà essere esplicitata agli utenti destinatari (manager) attraverso l'uso di schemi e grafici che diano un impatto visivo immediato dei risultati.

Si può notare come le prime fasi siano simili a quelle caratterizzanti l'architettura di un DW, infatti il data mining può essere considerato come l'evoluzione di un sistema OLAP integrato da analisi più approfondite e automatizzate⁵³.

⁵³ Cfr. PIGHIN M., MARZONA A., *Sistemi informativi aziendali, ERP e sistemi di data analysis*, Terza edizione, Pearson, 2018, pag. 321.

3.3 LE FUNZIONI DI MINING

Le attività di mining si distinguono in:

- *mining descrittivo*, tramite il quale vengono estratte informazioni relative alle proprietà generali dei dati (Pighin e Marzona, 2018); qui l'intento è quello di individuare gli schemi di regole presenti nel data warehouse che serviranno ad incrementare il livello di conoscenza. Affinché ciò si verifichi, le regole da identificare devono essere originali e non banali.
- *mining predittivo*, che a partire dai dati presenti ma soprattutto passati, determina regole generali e crea modelli da utilizzare per prevedere le tendenze future (Pighin e Marzona, 2018). La predizione è utilizzata per cercare di capire se e come il valore che assume una variabile casuale cambierà in futuro e con quale probabilità (in genere viene utilizzata questa tecnica per stimare le probabilità di abbandono dei clienti).

3.4 METODOLOGIE DI DATA MINING

I compiti principali per il data mining sono:

- **Descrizione:** è utile che i dati siano, almeno in una fase iniziale, descritti in maniera sintetica ed efficace. La descrizione dei dati può offrire importanti

suggerimenti in fase di analisi ed è utilizzata per sottolineare l'importanza delle relazioni e delle gerarchie tra i dati stessi.

- **Classificazione:** consiste nell'organizzazione dei dati in classi. La classe è un raggruppamento di elementi aventi le stesse caratteristiche e regole simili. In questa fase si parte da un campione di dati di cui è nota la classe di appartenenza, per poi ottenere le regole matematiche utili a descrivere il modello generale di classificazione.

La classificazione si rileva particolarmente utile come punto di partenza per effettuare analisi predittive, infatti è utile soprattutto per capire le caratteristiche dei clienti a rischio di abbandono.

- **Clusterizzazione** (o segmentazione): è simile alla precedente in quanto permette di raggruppare i dati in base alle similitudini ma, a differenza della classificazione, qui i dati non vengono raggruppati in base a regole e algoritmi matematici definiti dall'utente⁵⁴. Infatti, le regole di similitudine sono determinate dal sistema in base ai valori assunti dai dati, perciò in base agli elementi appartenenti ad ogni cluster si potranno ricavare

⁵⁴ A differenza della classificazione infatti, nel clustering non esistono classi predefinite che indicano l'appartenenza ad una classe di attributi ma si raggruppano semplicemente i dati in base all'omogeneità, a prescindere dalla classe di appartenenza.

ulteriori informazioni si di cui l'utente non era a conoscenza (Pighin e Marzona, 2018).

- **Associazione:** consiste nella scoperta di nessi causali ricorrenti estrapolabili dai dati⁵⁵. Le regole associative stabiliscono che se si verifica un evento A ci sarà una certa probabilità che si verifichi anche l'evento B.
- **Regressione:** anche questa è simile alla classificazione, dalla quale si differenzia per il fatto che le variabili (ossia le regole di appartenenza a una classe) nel caso delle regressioni possono assumere un numero elevato o infinito di valori⁵⁶.
- **Time series** (o serie storiche): insieme di dati riferiti al passato che si rivelano particolarmente utili per effettuare analisi di tipo predittivo e quindi cercare di capire quale potrà essere il comportamento di certi attributi in futuro.

Altra funzione è riferita al presente: si possono utilizzare serie storiche per descrivere i trend di certi fenomeni e sottolinearne la regolarità⁵⁷.

- **sequence discovery** (scoperta di sequenze): riprende il concetto di associazione ma applicando il fattore di correlazione sequenziale, ossia

⁵⁵ <http://www.intelligenzaartificiale.it/data-mining/>

⁵⁶ <http://www.intelligenzaartificiale.it/data-mining/>

⁵⁷ PIGHIN M., MARZONA A., *Sistemi informativi aziendali, ERP e sistemi di data analysis*, Terza edizione, Pearson, 2018, pag. 325.

rilevando quando al verificarsi di un evento A (esempio, acquisto di una casa), quasi sempre, *segue* l'evento B (acquisto in un certo arco temporale successivo di un frigorifero).

I settori di applicazione del data mining sono i più disparati, tra cui troviamo:

- marketing;
- economia e finanza;
- scienza;
- tecnologie dell'informazione e della comunicazione (ICT);
- statistica;
- industria⁵⁸.

In ambito di marketing relazionale, le principali applicazioni del data mining riguardano:

- **Clusterizzazione** della clientela: individuazione di tipologie di acquirenti che sono accomunati da abitudini di acquisto e caratteristiche sociodemografiche simili;

⁵⁸ Fonte: <http://www.intelligenzaartificiale.it/data-mining/>

- **Customer retriever**: analisi dei comportamenti dei clienti per identificare quali di questi sono a rischio abbandono e adottare quindi opportune strategie per impedirlo;
- **Market basket analysis** (analisi delle combinazioni di acquisto): si serve dell'analisi delle associazioni per capire quali prodotti o servizi vengono abitualmente comprati insieme;

3.5 GLI STRUMENTI DI MINING DEI DATI

Gli strumenti di data mining sono diversi e, a seconda dell'obiettivo da raggiungere e della difficoltà di analisi, potranno essere utilizzati in maniera differente oltre che complementare.

Tra i principali strumenti di data mining, troviamo le reti neurali, gli alberi di decisione e il Naïve Bayes (classificazione bayesiana).

Le **reti neurali artificiali** (RNA) sono strumenti di calcolo che funzionano in maniera simile al cervello umano. Infatti, le reti neurali sono definite tali perché costituite da *neuroni* (artificiali appunto) interconnessi da legami definiti *sinapsi*, al cui interno si trovano gli algoritmi di calcolo da applicare ai dati. Si caratterizzano quindi di un algoritmo di apprendimento che può essere di tipo lineare o non lineare e che gli permette, partendo dai campioni di dati, di ottenere i parametri

che rappresentano la conoscenza del problema analizzato per individuare una soluzione al problema stesso.

Una rete neurale è strutturata su 3 livelli (layer) di neuroni:

1. Layer formato dai neuroni di input, con cui vengono forniti i dati;
2. Uno o più layer intermedi in cui viene effettuata l'elaborazione dei dati;
3. Il layer di output che fornisce il risultato/soluzione.

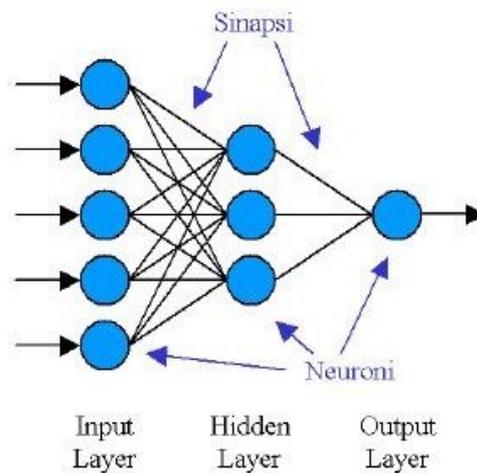


Figura 11. Reti neurali artificiali (Fonte: <http://www.mokabyte.it/2006/12/joone/>)

Le reti neurali possono funzionare secondo due diverse modalità: ad *apprendimento supervisionato o non supervisionato*.

Nelle prime, vengono inseriti i dati con le possibili soluzioni ai problemi e si programma un algoritmo in grado di fornire automaticamente le soluzioni corrette attraverso il rilevamento delle associazioni esistenti tra i dati.

Nelle reti neurali ad apprendimento non supervisionato invece, vengono inseriti solo alcuni dati di input selezionati e il sistema effettua le classificazioni attraverso il riconoscimento di analogie e differenze.

Le reti neurali costituiscono uno strumento di analisi dei dati molto flessibile dato che possono essere applicate in diversi ambiti e risultano particolarmente utili nell'analisi dei big data, ma hanno l'inconveniente di essere onerose oltre a richiedere molto tempo e conoscenze specifiche.

Altro strumento di mining è costituito dall'**albero di decisione** (decision tree), "un grafo in cui partendo dalla radice (training set), si procede ad effettuare una classificazione tramite un percorso che è di volta in volta una **scelta fra varie diramazioni**, o sottoinsiemi (detti nodi), i cui rami sono le alternative che conducono alle diverse foglie (risultati o classi)"⁵⁹. Costituiscono gli strumenti più utilizzati perché si adattano facilmente a qualunque tipo di dati, ma anche perché presentano un output molto chiaro che è rappresentato graficamente proprio come un albero (vedi figura 5).

Un albero di decisione efficace deve avere dimensioni adeguate, ossia non deve contenere troppe variabili altrimenti risulta essere caotico e di difficile comprensione ma soprattutto poco efficiente.

⁵⁹ <http://www.intelligenzaartificiale.it/data-mining/>

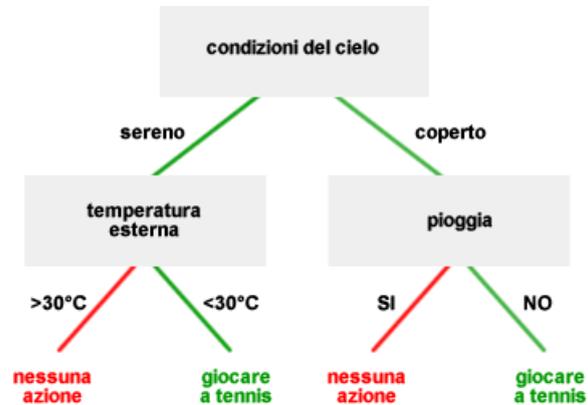


Figura 12. Esempio di albero decisionale (Fonte https://www.okpedia.it/albero_decisionale/).

Nel data mining, gli alberi di decisione sono utilizzati per operazioni di segmentazione, classificazione, regressione e serie storiche⁶⁰.

Il **Naive Bayes**, o classificazione bayesiana, è invece uno strumento che si basa sul Teorema di Bayes (da cui prende il nome), ossia sulla probabilità condizionata che un elemento appartenga o meno ad una determinata classe, in base agli attributi che caratterizzano l'elemento.

La particolarità di questo strumento è data dal fatto che ogni attributo viene valutato in modo indipendente rispetto agli altri: le probabilità di appartenere ad una classe non dipendono dalla relazione con gli altri attributi.

⁶⁰ <http://www.intelligenzaartificiale.it/data-mining/>

Questo strumento si rivela particolarmente utile quando la qualità dei dati è compromessa dalla presenza di errori, inconsistenze e dati anomali perché gli attributi risultano essere indipendenti ad altri e quindi non inficiano l'analisi di questi ultimi; il lato negativo sta invece proprio alla base del modello: è difficile che gli attributi siano tra loro indipendenti.

CAPITOLO 4: LA CONOSCENZA DEL “SOGGETTO BENEFICIANTE”: IL CASO “LEGA DEL FILO D’ORO ONLUS”

PREMESSA

Prima di entrare nel merito dell’analisi svolta, è opportuno presentare la Lega del Filo d’Oro, realtà in cui ho avuto occasione di collaborare tramite la partecipazione ad un progetto organizzato da “L.I.V.E. Intangibles”, Spin-off dell’Università Politecnica delle Marche dedicato alla Consulenza Manageriale.

Dopo una breve storia dell’associazione in cui si introduce all’attività che viene ancora oggi svolta, si descrive l’importanza attribuita alla Raccolta Fondi analizzando nello specifico le donazioni effettuate tramite i Lasciti Testamentari.

Segue l’analisi effettuata sul CRM e le modalità tramite le quali si è arrivati alla creazione del Datawarehouse, cercando di ottenere una certa conoscenza sul profilo del “beneficiante tipo” tramite il Data mining con Power BI.

Le fonti utilizzate per sviluppare tale capitolo sono costituite da Documenti che mi sono stati concessi direttamente dall’Associazione, quali “La luce dentro” (Biografia della fondatrice), Il Bilancio Sociale 2017, Il Bilancio di Sostenibilità 2018 e “I Libri della Lega del Filo d’Oro”.

4.1 LA STORIA DELLA LEGA DEL FILO D'ORO

La “Lega del Filo d’Oro” è un’Organizzazione Non Lucrativa di Utilità Sociale (ONLUS) nata nel 1964 come Associazione dalla volontà di Sabina Santilli, divenuta sordocieca a soli 7 anni a seguito di una meningite. La volontà di Sabina era quella di riunire tutti i sordociechi che conosceva, e anche quelli che non conosceva, per farli uscire dall’isolamento, metterli in contatto con il mondo esterno e renderle persone per quanto possibile autonome.

La sua voglia di emergere l’ha contraddistinta fin da piccola. Nel 1927 all’età di 10 anni diventa la prima alunna a frequentare a Roma l’Istituto per educatori non vedenti “Regina Margherita di Savoia”, anno in cui l’educatore Augusto Romagnoli (da cui prende nome l’Istituto oggi) ha iniziato ad adottare il suo metodo educativo per ciechi essendo affetto lui stesso da tale disabilità. A 22 anni Sabina si iscrive all’Istituto Professionale per Ciechi di Firenze perché voleva “imparare il mestiere” per guadagnarsi da vivere da sola. Tornata in Abruzzo nel suo paese d’origine, a 31 anni riprese tutte le sue attività, cercando di non risultare mai come un “peso” per la sua famiglia bensì un aiuto. Si dava da fare nelle faccende domestiche come aveva sempre fatto, chiese la modifica di un pezzo della macchina da cucire alla Necchi per evitare di farsi male e fece tradurre il manuale di taglio e cucito in Braille oltre che per lei, per tutte le donne che come lei volevano continuare a

cucire nonostante la mancanza di uno dei sensi che più serviva per svolgere tale attività, quale la vista. Il suo spirito, la volontà di farcela sempre e comunque l'ha spinta ad attivarsi per coloro che erano nella sua stessa situazione e venivano considerati *stupidi, malati di mente, irrecuperabili, diversi*.

Era una donna a dir poco straordinaria, verrebbe da pensare che fosse una donna d'altri tempi, anche se alcuni la definivano proprio donna del suo tempo, donna tipica degli anni Cinquanta che studiava, leggeva, ricamava, cuciva.

Nel gennaio del 1964 incontra Don Dino Marabini, sacerdote di Osimo allora Direttore Nazionale del Movimento Apostolico Ciechi, per ricevere la sua assistenza. Fu lui ad abbracciare l'idea di Sabina di creare un'associazione che unisca i sordociechi e li aiuti nelle attività quotidiane.

L'associazione nasce il 20 dicembre 1964 a San Benedetto dei Marsi (AQ), paese di origine di Sabina, insieme a Don Dino Marabini e 20 volontari e pochi anni dopo si sposta ad Osimo nel 1967. Sabina fu così la prima donna sordocieca in Italia a stipulare un atto legale e ad assumere una carica sociale, ossia quella di Presidente della sua associazione.

“Tale associazione è nata per la ferrea volontà e per il deciso impegno di una donna che conosce molto bene i ciecosordi, la loro vita, il loro stato di abbandono, le loro condizioni nei ricoveri non adatti alla loro esistenza, raccogliendo giorno per

giorno, le loro ansie e le loro pene”⁶¹ queste le parole di Loda Santilli, sorella della fondatrice. Solo una persona come Sabina poteva pensare al modo in cui poter rimediare o migliorare tali condizioni perché poteva capire benissimo lo stato d’animo di queste persone, nonostante la fortuna di aver avuto un forte appoggio da parte della sua famiglia, cosa molto difficile a quel tempo in cui invece le famiglie non potevano che essere rassegnate al loro destino.

Il nome dell’associazione, abbastanza particolare, è dovuto alla metafora di Sabina secondo cui vi è un “filo aureo” di amicizia che unisce i sordociechi tra loro e con coloro che sono disposti ad aiutarli per conseguire una certa autonomia e vivere una vita al pari degli altri.

Fin dalla sua fondazione, la Lega del Filo d’Oro offre assistenza, riabilitazione e reinserimento nella famiglia e nella società delle persone sordocieche e pluriminorate psicosensoriali seguendo gli stessi principi della fondatrice ma anche introducendo nuove modalità educative per rispondere sempre meglio alle diverse esigenze che si presentano, rappresentando un valido punto di riferimento per le persone affette da sordocecità, a cui spesso sono legate altre disabilità (motorie, intellettive, malformazioni scheletriche o dentali...). Inoltre, offre un valido supporto per le famiglie dei suoi utenti, perché la stessa Sabina aveva

⁶¹ SANTILLI L., *La luce dentro e altri scritti su Sabina Santilli*, Cooperativa Sociale CO.A.L.A Onlus, Jesi (AN), 2003.

intuito come la *comprensione* dei familiari e la loro collaborazione costituiscano un valido aiuto per le persone sordocieche; alle famiglie viene offerto un supporto psicologico, lavori di gruppo con altri genitori, laboratori e il cosiddetto “parent training” (formazione dei genitori e dei familiari) che favorisce i percorsi riabilitativi e il reinserimento nella famiglia.

Quando fu fondata, le prime attività dell’associazione consistevano in soggiorni estivi organizzati da Sabina per spingere giovani adulti sordociechi a fare amicizia e uscire dall’isolamento. Oggi invece, grazie alla ricerca e all’impegno dell’associazione si è arrivati a costituire un vero e proprio centro di eccellenza dedicato per tutte le età. Negli ultimi anni si è investito molto sull’intervento precoce, che riguarda i bambini da 0 a 4 anni, per cercare di intervenire tempestivamente, fin dalla nascita, per evitare peggioramenti della situazione di disabilità. Oggi i bambini sottoposti a trattamenti da parte della lega del filo d’oro costituiscono il 35% degli utenti totali e la maggior parte di questi ha bisogno di un intervento precoce.

L’associazione è oggi presente in Italia con 8 sedi di cui 3 territoriali (Padova, Roma e Napoli) e 5 sedi Riabilitativo – Residenziali (Osimo, Lesmo, Molfetta, Modena, Termini Imerese) tra cui figura appunto la sede centrale di Osimo (AN). Attualmente sono in fase di apertura due ulteriori sedi a Novara e Pisa.

La sottile differenza tra le sedi Riabilitativo – Residenziali e quelle territoriali sta nel fatto che queste ultime si limitano ad offrire servizi di assistenza, riabilitazione, servizi ambulatoriali, ricreativi e di supporto alle famiglie tramite l'integrazione rispetto ai servizi offerti sul territorio, anche con la collaborazione di altre associazioni, mentre le prime offrono assistenza a 360 gradi e, nei casi più complessi, possono offrire ricovero per gli utenti che ne hanno bisogno e alloggio per le rispettive famiglie.



Figura 13. Esterno del Nuovo Centro Nazionale di Osimo (Fonte: https://www.repubblica.it/solidarieta/disabilita/2019/05/15/news/a_osimo_il_nuovo_centro_della_lega_d_el_filo_d_oro-226351987/)

Dall'ultimo Bilancio di Sostenibilità (2018) emerge che gli utenti che hanno usufruito di uno o più servizi presso le diverse sedi sono aumentati da 892 (del

2017) a 911 e di questi, 611 sono stati seguiti tramite trattamenti intensivi o estensivi, con un livello di soddisfazione medio delle loro famiglie pari al 90%.

Le prestazioni effettuate dalla Lega del Filo d'oro sono fruibili gratuitamente per i propri utenti, e ciò è reso possibile grazie ad una sovvenzione con il Servizio Sanitario Nazionale e ai finanziamenti pubblici che permettono all'associazione di ricevere delle rette giornaliere ma solo per i ricoveri e le prestazioni ambulatoriali, corrisposte dalle ASL di provenienza dei propri utenti. È importante però sottolineare che le rette non risultano essere sufficienti a coprire tutti i costi sostenuti per i servizi offerti, ma permettono una copertura di circa il 33% di questi in virtù dell'elevato livello di servizio erogato agli utenti.

Le risorse che permettono di erogare servizi di assistenza, cura, riabilitazione, ma anche effettuare ricerca scientifica ed innovazione derivano principalmente dalle attività di raccolta fondi con le quali si riesce coprire circa il 67 % dei costi sostenuti, motivo per il quale risulta essere un'attività estremamente fondamentale e su cui si investe molto in termini di comunicazione e coinvolgimento dei sostenitori.

4.2 L'IMPORTANZA DELLA RACCOLTA FONDI

Uno degli obiettivi che l'organizzazione si pone è quella di consolidare i rapporti e le relazioni con i sostenitori, tramite le campagne di sensibilizzazione per le

donazioni regolari e sporadiche, 5 per mille e lasciti; fondamentale inoltre è la creazione di un legame forte con le istituzioni e la pubblica amministrazione.

Molte sono le iniziative a supporto dell'attività di Raccolta Fondi; tra queste risultano significative le iniziative di Direct Marketing rivolte sia agli attuali sostenitori che ai donatori potenziali, per citarne alcune: la Giornata del Sostenitore, la Campagna SMS solidale, il Face to Face e il Direct Response TV (questi ultimi due mirati a promuovere le donazioni regolari).

Tramite queste attività si sono ottenuti nel corso degli anni importanti risultati in termini di aumento delle donazioni, nonostante una lieve contrazione del numero di donatori.

Le risorse provengono principalmente da donatori "one off" e dalle elargizioni in termini di Eredità e Lasciti testamentari, buona parte deriva anche dal 5 per mille ed in misura minore da donatori regolari e attività digitali (Figura 13).

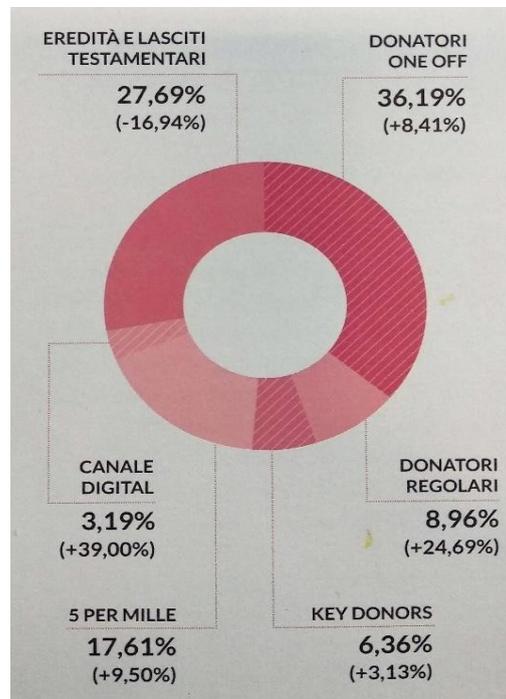


Figura 14 Ripartizione delle donazioni per tipologia di canale (variazione rispetto al 2017). (Fonte: Bilancio di Sostenibilità 2018)

Come si nota dal grafico, le risorse pervenute in termini di eredità e lasciti rappresentano quasi il 28% delle donazioni totali, rappresentando la seconda principale risorsa dopo le donazioni sporadiche (one off).

Sono anche da considerare l'attività di merchandising (bomboniere e regali solidali), la raccolta fondi tramite Facebook e il rapporto con aziende partner, enti e fondazioni che permettono di sostenere progetti specifici come, ad esempio, la realizzazione del secondo lotto del nuovo Centro Nazionale.

4.3 IL TESTAMENTO SOLIDALE

Uno dei punti di forza della Lega del Filo d'Oro è l'**attività della raccolta** fondi, ottenuta grazie alla creazione di relazioni con i sostenitori e i partner.

In quest'ambito, risultano da tempo fondamentali anche le iniziative rivolte al fine di ottenere risorse da eredità e lasciti testamentari, che rappresentano come visto una risorsa abbastanza importante.

Quando arriva il momento di disporre delle ultime volontà e metterle per iscritto in un testamento, sempre più italiani decidono di fare del bene e donare una parte del loro patrimonio in beneficenza, ad associazioni umanitarie, che si impegnano nel sociale o nella ricerca scientifica, più in generale, agli enti del Terzo Settore (associazioni, Onlus...). In questi casi si parla di Testamenti Solidali o Lasciti Testamentari.

“Fare “testamento solidale”, significa ricordare nel proprio testamento, in qualità di erede (eredità) o di legatario (legato), una o più associazioni, organizzazioni, enti”⁶².

Se il destinatario del testamento è un *erede*, avrà diritto a ricevere tutto o parte del patrimonio mobiliare e immobiliare di colui che effettua il testamento, mentre

⁶² Per ulteriori approfondimenti: <https://www.notariato.it/it/tags/testamento-solidale>

nel caso in cui un soggetto viene considerato *legatario* avrà diritto al solo bene specifico che gli viene attribuito all'interno del testamento.

Il testamento può essere redatto in forma **pubblica** presso un notaio alla presenza dei testimoni oppure scritto a mano dallo stesso testatore presso il proprio domicilio (si parla in questo caso di testamento di tipo **olografo**) oppure si può decidere di scriverlo in **forma segreta** da consegnare comunque al notaio in busta chiusa con sigillo.

“È un gesto semplice e non vincolante, che può essere ripensato, modificato in qualsiasi momento, senza che vengano in alcun modo lesi i diritti successori dei propri cari e familiari che sono tutelati dalla legge italiana con il riconoscimento della quota di legittima, ovvero la parte di eredità di cui si può disporre liberamente individuando come beneficiari soggetti diversi dai legittimari, come le onlus. Non sono necessari ingenti patrimoni, perché per sostenere il lavoro quotidiano di associazioni impegnate nelle più importanti cause umanitarie e scientifiche, anche un piccolo contributo può fare la differenza”⁶³.

In altri Paesi come ad esempio negli Stati Uniti, fare testamento a favore di tali organizzazioni è un gesto più che diffuso in quanto l'ordinamento giuridico non riconosce tutele per gli eredi legittimi, ossia per i familiari più stretti, per cui non

⁶³ <https://www.notariato.it/it/tags/testamento-solidale>

ci si sente in dovere di destinare una parte di patrimonio a questi ultimi. Proprio per questi motivi negli USA si registra una forte propensione alla destinazione dei propri averi in beneficenza, soprattutto da parte di personaggi famosi come Sting, Bill Gates, Robin Williams che hanno destinato la maggior parte dei propri beni a favore di tali enti.

Rossano Bartoli, Presidente della Lega del Filo d'Oro nonché portavoce del Comitato Testamento Solidale, sostiene che «Per fare del bene non bisogna disporre di ingenti patrimoni o essere un personaggio famoso, anche un piccolo gesto può aiutare concretamente le Organizzazioni del Comitato a portare avanti i loro progetti in Italia e nel mondo. Siamo orgogliosi degli importanti risultati ottenuti finora, ma il nostro lavoro non si ferma qui, anzi [...] dobbiamo porci nuovi obiettivi per far conoscere sempre più lo strumento del testamento solidale». Non è detto quindi che le somme da donare debbano essere necessariamente esorbitanti anzi, basta un piccolo gesto per fare “qualcosa di grande”.

Nonostante gli ultimi anni in Italia si sia registrata una riduzione del numero delle donazioni e dei donatori, si è verificata invece una crescita delle donazioni tramite un lascito testamentario. Infatti, secondo un articolo pubblicato da Vita del 12 settembre 2019 sono sempre più gli italiani che vengono a conoscenza della possibilità di poter contribuire all'erogazione di servizi da parte delle organizzazioni non lucrative, lasciando loro un ricordo quando non ci saranno più.

La conoscenza di tale strumento è cresciuta anche grazie alla nascita del Comitato Testamento Solidale nel 2013, a cui partecipano attualmente 21 organizzazioni no profit: ActionAid, AIL, AISM, Fondazione Don Gnocchi, Lega del Filo d'Oro, Save the Children, Aiuto alla Chiesa che Soffre onlus, Amnesty International, AMREF, CBM, CESVI, Greenpeace, Intersos, Fondazione Operation Smile Italia onlus, Fondazione Pasteur, Fondazione Telethon, Fondazione Umberto Veronesi, Progetto Arca, Telefono Azzurro, Unicef, Università Campus Bio-Medico di Roma, a cui si aggiunge la partecipazione del Consiglio Nazionale del Notariato, che offre supporto informativo circa le tipologie di testamento e relative modalità con le quali poter inserire le associazioni beneficiarie e sottolinea quali sono le tutele sia a favore degli eredi legittimi⁶⁴ che del testatore stesso, il quale può in qualsiasi momento modificare, annullare o integrare le disposizioni precedentemente indicate.

Si pensa quindi che la crescita del numero di testamenti di tipo solidale è dovuta al fatto che in alcuni casi sono gli stessi notai ad informare i testatori dell'esistenza di tale possibilità, portando ad una maggiore conoscenza di tale strumento per i testatori, ma non è necessariamente detto.

⁶⁴ Gli eredi legittimi sono coloro che in automatico subentrano nella successione, salvo diverse disposizioni, a prescindere dalle volontà del testatore. In questa categoria di soggetti rientrano il coniuge, i figli, i genitori, fratelli, sorelle e i parenti entro il sesto grado.

4.4 L'INTRODUZIONE DEL CRM

Al fine di rafforzare la relazione con i sostenitori, che costituiscono elemento imprescindibile, la Lega del Filo d'Oro ha deciso di introdurre recentemente il CRM per la gestione dei dati e dei processi all'interno dell'organizzazione, in occasione del trasferimento di tutte le attività presso il Nuovo Centro Nazionale di Osimo nel 2018.

L'obiettivo a cui si voleva arrivare era l'introduzione di un sistema adatto alla raccolta fondi e in particolare all'attività di comunicazione nell'ottica di considerare il donatore al centro dell'attività di marketing relazionale, ottenendo una migliore profilazione e personalizzando le comunicazioni.

Prima dell'introduzione del CRM i dati prodotti e/o rilevati nelle varie attività dell'associazione venivano inseriti all'interno di fogli Excel e in Database Access che erano gestiti in autonomia da ogni ufficio, perciò si trattava di file di tipo tematico strettamente legati all'attività svolta dall'ufficio stesso e non collegati in alcun modo alle altre attività.

I **primi database in formato Access** furono creati nel 2009 internamente, dal personale dedicato ai Servizi Informatici per i diversi uffici, i quali avevano così ognuno il proprio database da alimentare e consultare. In quel periodo inoltre gli

uffici erano dislocati in varie sedi (anche quelli di Osimo) per cui la mancanza di comunicazione era molto sentita.

I database erano un po' scarni, caratterizzati da pochi campi oltre a quelli essenziali e come ancora oggi accade nel nuovo CRM, non sempre completati.

Grazie all'introduzione del CRM sono stati inseriti ulteriori campi che hanno permesso di poter effettuare una profilazione più accurata dei sostenitori, attraverso l'aggiunta di una serie di campi, in particolare nell'ambito comunicazione per fare in modo che ogni sostenitore riceva le informazioni tramite i canali che più preferisce (corrispondenza, email e così via).

A seguito di un periodo di crescita del numero delle donazioni e dei soggetti donanti dovuto anche all'aumento delle donazioni on-line, nel 2016 nasce l'esigenza da parte dell'ufficio di Raccolta Fondi di dotarsi di un sistema informativo che fosse più veloce, flessibile, in grado di gestire una quantità di dati decisamente superiore a quella che potevano contenere i database Access. In quell'anno l'ufficio Raccolta fondi (che tratta più dati in assoluto) registrava circa 7 milioni di record⁶⁵ sulle donazioni e 1,5 milioni di record di anagrafiche inserite. A partire quindi dal 2016 si inizia ad effettuare un'analisi interna riguardo le esigenze conoscitive dei diversi uffici, cercando di capire quali erano i dati *attuali*

⁶⁵ Un **record** rappresenta tutti i dati (contenuti nei vari campi) di una singola voce nella tabella di un database (<https://www.mrwebmaster.it/sql/cosa-record-8202.html>).

di maggiore interesse presenti all'interno dei database già presenti e ancora validi da considerare, e soprattutto i nuovi dati da trattare anche in vista degli obiettivi futuri da raggiungere. Contemporaneamente a quest'analisi si decide di effettuare un confronto con altre associazioni nazionali ed internazionali già dotate di software di questo tipo, per cercare di capire come questo poteva essere strutturato e quali variabili potevano essere introdotte nel nuovo CRM.

Il passo successivo è stato quello di individuare il fornitore di software più adatto alle esigenze dell'organizzazione e dopo averne esclusi diversi, la scelta dovette avvenire tra due soluzioni: una dedicata strettamente alle Organizzazioni Non-profit ma avente una struttura rigida, quindi difficilmente adattabile alle esigenze, l'altra dedicata alle Aziende profit ma con ampie possibilità di modifica, adattamento e personalizzazione.

La scelta è ricaduta su quest'ultimo, ossia Microsoft Dynamics 365 nonostante abbia comportato non poche sfide per l'organizzazione, dovute sia ai continui adeguamenti richiesti dagli uffici coinvolti (inizialmente solo Comunicazione - Raccolta fondi e Ufficio Affari generali), sia alla trasposizione dei dati dai database precedenti al nuovo, e alla formazione del personale dipendente che fa uso del programma quotidianamente.

Il 2017 è stato l'anno di transizione in cui si è effettuata a scaglioni la migrazione dei dati presenti nei database Access all'interno del CRM. Il passaggio è avvenuto

in maniera graduale, esportando i dati prima in tabelle Excel e poi nel sistema, effettuando continue verifiche a campione.

Il CRM, come anzi detto, diventa operativo a tutti gli effetti nel 2018, con lo spostamento nonché l'accentramento di tutte le attività d'ufficio presso il Centro Nazionale di Osimo ma solo per le attività di Comunicazione, Raccolta fondi e per gli Affari Generali.

4.5 IL PROGETTO “UN RICORDO TANGIBILE”

Da tempo la Direzione aveva l'obiettivo di intraprendere un percorso di analisi delle pratiche di Lascito testamentario pervenute negli anni in quanto rappresentano una risorsa particolarmente strategica, un punto di forza dell'attività di raccolta fondi della Lega del Filo d'Oro.

Alla luce della disponibilità dei dati già inseriti in CRM che inizialmente si presentavano abbastanza incompleti e date le innumerevoli possibilità che il programma offre, si è deciso di approfondire la conoscenza su questo tema avviando il progetto “Un ricordo tangibile” con la collaborazione di “L.I.V.E. Intangibles”.

L'ambito di interesse è stato quello dell'analisi del **CRM** per gestire in chiave strategica i dati riguardanti i testamenti solidali ricevuti.

Lo scopo di questo elaborato è uno studio sulla struttura ed il funzionamento dei dati esistenti con l'obiettivo di poter svolgere un'analisi di Business intelligence ed ottenere informazioni utili ad effettuare previsioni e iniziative mirate.

L'analisi si è concentrata in questa prima fase limitatamente alle pratiche di lascito testamentario effettuate a favore della Lega del Filo d'oro a partire dal 2014 e fino al 2019, analizzando le specificità delle pratiche pervenute ma principalmente le caratteristiche del "donatore tipo" cercando di suddividere i donatori in cluster, quindi in base alle similitudini.

La limitazione temporale dell'analisi deriva dal fatto che era chiara a tutti la mancanza di dati utili nel CRM e la bassa qualità dei dati già presenti, perciò si è reso necessario dapprima bonificare la base dati, selezionare i dati più utili per produrre le informazioni richieste e poi dedurre delle considerazioni.

La prima fase del progetto ha visto lo studio di "**Microsoft Dynamics 365**" il CRM introdotto nel 2018 per la Raccolta Fondi.

Una volta appreso il funzionamento del sistema, si sono analizzati tutti i campi riguardanti l'Anagrafica dei sostenitori e la Pratica Legale riferita al lascito.

Il CRM si presentava all'apparenza abbastanza "completo" in quanto caratterizzato da moltissimi campi potenzialmente utili all'analisi, anche se molti purtroppo erano non compilati.

Da qui, la necessità di **analizzare i campi** di maggiore interesse al fine di capire quali elementi potevano essere utili/meno utili, interessanti e individuabili. In questa fase, è risultato necessario il confronto continuo con Francesca Silvestri, Project Leader dell'introduzione del CRM, con cui si è effettuata una **selezione** di campi utili ai fini analitici e valutata l'**integrazione** di qualche campo aggiuntivo. Prima di definire i campi "chiave" abbiamo ritenuto opportuno leggere alcuni testamenti.

Dalla prima lettura dei testamenti è emerso che sostanzialmente la struttura del documento è abbastanza omogenea, soprattutto per chi lo sottoscrive alla presenza di un notaio; questo però non garantisce che le informazioni ricercate siano sempre presenti, data l'impronta personale che lo caratterizza. Trattandosi di un documento strettamente personale infatti, non sempre sono presenti i dati di cui si vorrebbe disporre.

Al fine di contenere gli sforzi di "bonifica" dei dati abbiamo quindi scelto i campi da migliorare, sulla base del principio costi/benefici.

Abbiamo quindi costruito il modello di datawarehouse che poteva rappresentare il punto di partenza dell'analisi di Business Intelligence, mostrando le relazioni fondamentali tra tabelle delle dimensioni e le tabelle dei fatti, formando uno schema inizialmente abbastanza semplice, seppur costituito da diverse tabelle.

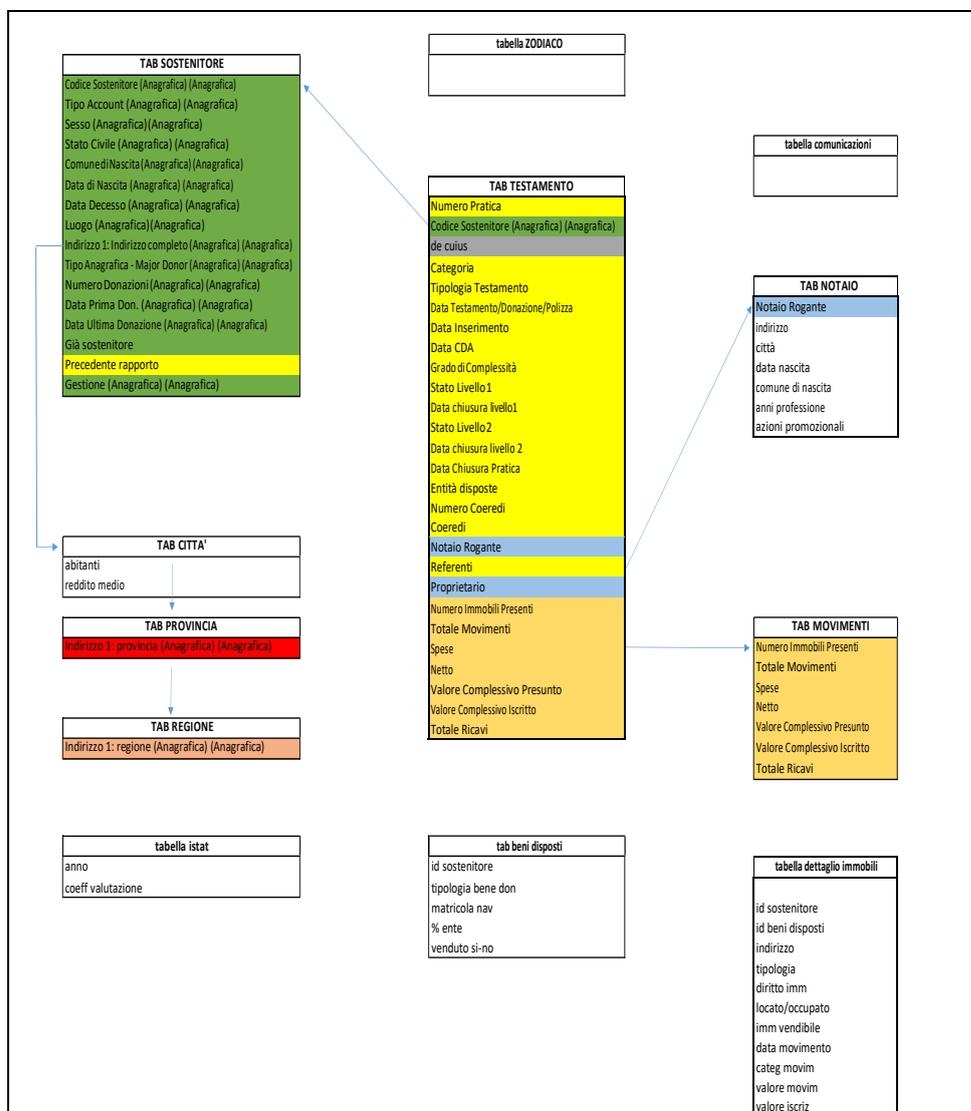


Figura 15. Il primo tentativo di creazione di uno schema di Data Warehouse.

La “tabella testamento” rappresenta la tabella del fatto, a cui sono legate le tabelle delle dimensioni di analisi che inizialmente si pensava di considerare, quali i dati del sostenitore, la provenienza (suddivisa gerarchicamente in Regione, Provincia,

Città), le comunicazioni inviate, il notaio, i movimenti contabilizzati riferiti alla pratica, i beni mobili e immobili, il segno zodiacale (per evidenziare le relazioni tra i lasciti ed il segno zodiacale) e la tabella Istat contenente il reddito dei comuni.

Nonostante le innumerevoli verifiche effettuate a campione sui dati trasferiti nel nuovo CRM, i dati non si sono rivelati sempre corretti ma anzi si presentavano alcune volte errati e spesso di scarsa qualità perché contenuti all'interno di campi prettamente descrittivi oppure non compilati.

Molte volte gli errori sono stati individuati proprio nella fase di Data entry effettuata "caso per caso" in CRM e non sono stati rari i casi in cui gli errori e le mancanze sono stati individuati tramite il ricorso alle Tabelle Pivot create in Excel a partire dai dati estratti proprio dal CRM.

Per tutti questi motivi si è deciso di creare uno schema esplicativo delle problematiche riscontrate riguardo l'inserimento dei dati.

CAMPI	VISUALIZZAZIONE	TABELLA DEL DW	QUALITA' DEL DATO	NOTE QUALITA'	CAMPI MANCANTI	RESPONSABILE	NOTE	PROBABILITA' DI RISOLUZIONE
Numero Pratica	pratica legale	tab testamento	ALTA	num. progressivo	NO			già risolto
Codice Sostenitore (Anagrafica) (Anagrafica)	anagrafica	Tab sostenitore	ALTA	default	NO			già risolto
de cuius	pratica legale	tab sostenitore	ALTA	descrittivo	NO			già risolto
Tipo Account (Anagrafica) (Anagrafica)	anagrafica	Tab sostenitore	ALTA	problematica genere m/f	NO			ALTA
Sesso (Anagrafica) (Anagrafica)	anagrafica	Tab sostenitore	MEDIA	errata associazione sesso/sostenitore	NO			ALTA
Stato Civile (Anagrafica) (Anagrafica)	anagrafica	tab sostenitore	BASSA	compilato solo nelle mie	NO			ALTA
Comune di Nascita (Anagrafica) (Anagrafica)	anagrafica	Tab sostenitore	BASSA	rivedere e compilare	NO			ALTA
Data di Nascita (Anagrafica) (Anagrafica)	anagrafica	Tab sostenitore	ALTA	si ricava dal testamento	NO			già risolto
Data Decesso (Anagrafica) (Anagrafica)	anagrafica	Tab sostenitore	ALTA	campo data	NO			già risolto
Luogo (Anagrafica) (Anagrafica)	anagrafica	Tab sostenitore	MEDIA/BASSA	luogo decesso, campo libero	NO			ALTA
Indirizzo 1: indirizzo completo (Anagrafica) (Anagrafica)	anagrafica	tab indirizzo	MEDIA	Non sempre normalizzato	NO			ALTA
Indirizzo 1: regione (Anagrafica) (Anagrafica)	anagrafica	tab indirizzo	ALTA	elimina duplicati	NO			ALTA
Indirizzo 1: provincia (Anagrafica) (Anagrafica)	anagrafica	tab indirizzo	ALTA	elimina duplicati	NO			ALTA
Tipo Anagrafica - Major Donor (Anagrafica) (Anagrafica)	anagrafica	Tab sostenitore	BASSA	gestito solo dal 2018	NO			BASSA
Numero Donazioni (Anagrafica) (Anagrafica)	anagrafica	Tab sostenitore	ALTA	campo calcolato	NO			già risolto
Data Prima Don. (Anagrafica) (Anagrafica)	anagrafica	Tab sostenitore	ALTA	campo calcolato	NO			già risolto
Data Ultima Donazione (Anagrafica) (Anagrafica)	anagrafica	Tab sostenitore	ALTA	campo calcolato	NO			già risolto
Già sostenitore	pratica legale	tab sostenitore	ALTA	sempre specificato	NO			già risolto
Precedente rapporto	pratica legale	tab testamento	ALTA	sempre specificato	NO		rapporto informativo	già risolto
Gestione (Anagrafica) (Anagrafica)	anagrafica	Tab sostenitore	ALTA	sempre specificato	NO		campo anagrafica perché esiste a prescindere dall'esistenza di un testamento	già risolto
Categoria	pratica legale	tab testamento	ALTA	sempre specificato	NO			già risolto
Tipologia Testamento	pratica legale	tab testamento	MEDIA	compilato solo nelle mie	NO			ALTA
Data Testamento/Donazione/Polizza	pratica legale	tab testamento	ALTA	sempre inserito	NO			già risolto
Data Inserimento	pratica legale	tab testamento	ALTA	default	NO			già risolto
Data CDA	pratica legale	tab testamento	ALTA	sempre inserito	NO			già risolto
Grado di Complessità	pratica legale	tab testamento	MEDIA	non specificato per le vecchie	NO			BASSA
Stato Livello 1	pratica legale	tab testamento	ALTA	sempre specificato	NO			già risolto
Data chiusura livello1	pratica legale	tab testamento	ALTA	dipende se è stata chiusa o m	NO			già risolto
Stato Livello 2	pratica legale	tab testamento	MEDIA	dipende dalla complessità	NO			già risolto
Data chiusura livello 2	pratica legale	tab testamento	BASSA	sostituito da data ch pratica	NO			ALTA
Data Chiusura Pratica	pratica legale	tab testamento	MEDIA	manca nelle vecchie pratiche	NO			già risolto
Entità disposte	pratica legale	tab testamento	BASSA	campo libero descrittivo	NO			BASSA
Numero Coeredi	pratica legale	tab testamento	BASSA	non sempre specificato	NO			MEDIA (lettura testamento)
Coeredi	pratica legale	tab testamento	BASSA	campo libero	NO			BASSA
Notaio Rogante	pratica legale	tab notaio	MEDIA	895 vuoti (vecchie pratiche)	NO			MEDIA (lettura testamento)
Referenti	pratica legale	tab testamento	BASSA	campo libero	NO			MEDIA
Proprietario	pratica legale	tab testamento	MEDIA	848 non ricollegabili ad un no	NO			BASSA
Numero Immobili Presenti	pratica legale	tab movimenti	MEDIA (VEDI PIVOT)	campi vuoti e/o zero	NO			ALTA
Totale Movimenti	pratica legale	tab movimenti	ALTA		NO			già risolto
Spese	pratica legale	tab movimenti	ALTA		NO			già risolto
Netto	pratica legale	tab movimenti	ALTA		NO			già risolto
Valore Complessivo Presunto	pratica legale	tab movimenti	BASSA	spesso non compilato	NO			BASSA
Valore Complessivo Iscritto	pratica legale	tab movimenti	ALTA	solo se pervenuto accredito	NO			già risolto
Totale Ricavi	pratica legale	tab movimenti	ALTA	solo se pervenuto accredito	NO			già risolto
tipologia coeredi (campo codificato)	pratica legale	tab testamento	ALTA (potenziale)	da creare	SI		da creare	ALTA
Data Creazione	pratica legale	tab testamento	MEDIA	data import nuovo sistema	NO			BASSA
Nome account (Anagrafica) (Anagrafica)	anagrafica	tab sostenitore	ALTA	ridondante	NO		coincide con il de cuius	già risolto
Professione	anagrafica	tab sostenitore	BASSA	difficile reperimento	NO			BASSA
Età	anagrafica	tab sostenitore	MEDIA	campo calcolato	NO		adatto all'anagrafica dei soggetti in vita	MEDIA (calcolo in excel)

Figura 16. Tabella esplicativa delle problematiche connesse ai campi (nostra elaborazione).

Si può notare che ogni campo è stato analizzato singolarmente, inquadrandolo nella sezione in cui viene visualizzato e nella tabella del data warehouse, è stata rilevata la qualità del dato, la mancanza del campo, il responsabile, un campo destinato ad eventuali annotazioni e la probabilità di risoluzione del problema relativo a quel campo.

Questo schema ci ha consentito di individuare i campi “vitali” per la successiva analisi diagnostico/descrittiva ed ha permesso di effettuare un’ulteriore selezione attraverso l’eliminazione dei campi i cui dati erano difficilmente reperibili.

Per compilare la tabella in riferimento alla qualità dei dati abbiamo creato delle tabelle Pivot per capire quali erano i dati che presentavano le maggiori criticità. In particolare, vi erano campi in cui mancava una percentuale significativa di dati che col tempo non erano più stati compilati o che mancavano di una chiara tassonomia in quanto destrutturati.

Molti degli errori riscontrati, sono stati provocati dalla trasposizione dei dati che ha provocato: l'esistenza di date di nascita posteriori alle date di morte, l'inserimento solo parziale dei valori economici senza considerare le spese sostenute, la mancata trasposizione degli indirizzi o il trasferimento ma senza direttamente procedere alla Normalizzazione Google che permette il collegamento diretto con Maps, o ancora la mancata dichiarazione del Decesso. Quest'ultima risulta essere fondamentale per i report sui sostenitori attivi che venivano in questo modo inficiati dai dati su soggetti che non effettuavano più donazioni; in molti casi la data di nascita era inserita di default come "1900".

Molto spesso si è anche riscontrata la presenza di anagrafiche duplicate che fortunatamente possono essere integrate a formarne una sola; si sono inoltre incontrate un paio di casistiche in cui il soggetto beneficiante non era propriamente inserito nell'anagrafica ma al suo posto era registrato in anagrafica il soggetto che aveva effettuato il versamento per suo conto.

Rimane ancora da chiarire la possibilità di poter legare ad uno stesso soggetto più di una pratica perché magari una viene effettuata tramite Donazione con Atto Pubblico *inter vivos* ed un'altra in fase di redazione di testamento e che produrrà effetti dopo la morte.

In accordo con il team di progetto⁶⁶ si è deciso di effettuare l'analisi a partire dalle pratiche più recenti (anno 2019) che al momento di inizio della collaborazione erano già state approvate dal Consiglio di Amministrazione procedendo poi a ritroso, fino alle pratiche pervenute nell'anno 2014.

Data la recente introduzione del CRM attualmente in uso, le pratiche meno recenti mancavano anche degli allegati necessari all'analisi⁶⁷, motivo per il quale abbiamo dovuto proseguire anche con l'inserimento degli allegati presso l'Area Affari generali – Ufficio legale. In questo contesto ci è stato chiesto di allegare alla pratica, oltre che i testamenti mancanti, anche i documenti ufficiali più importanti (inventario e accettazione di eredità, dichiarazione di successione dell'agenzia delle entrate e le visure catastali con relative perizie ove fossero presenti immobili)

⁶⁶ Hanno partecipato al gruppo di lavoro: il Project Leader dell'introduzione del CRM, il responsabile dell'Ufficio Affari generali e la responsabile dell'area Amministrazione, Finanza e Controllo, oltre al Prof. Danilo Scarponi Consulente "L.I.V.E. Intangibles" e la sottoscritta.

⁶⁷ Per ogni pratica vengono allegati una serie di documenti di natura legale che contengono dati importanti anche ai fini della nostra analisi. Il principale allegato è il testamento a cui solitamente si aggiungono la dichiarazione di successione, accettazione e inventario di eredità, e altri.

affinché si potesse avere una panoramica generale ed allo stesso tempo completa della singola pratica e renderne più veloce la navigazione.

Attraverso la lettura dei testamenti abbiamo trovato quasi sempre le informazioni utili a popolare il CRM e, allo stesso tempo, abbiamo proposto l'introduzione di altri due campi quali la *tipologia di testamento* e lo *stato civile* che seppur non presenti nel sistema, sono apparsi interessanti sia per la nostra analisi che per altre future analisi interne. Altro campo d'interesse poteva essere la "professione" ma purtroppo questo non è sempre compilato perché si tratta di una caratteristica quasi mai specificata nei testamenti (a volte si possono fare delle supposizioni sulla base di alcuni dettagli rilevati in fase di lettura) anche nel caso dei Major Donor, nonostante questi ultimi condividono con l'associazione un maggior numero di informazioni.

Alcune volte abbiamo potuto ottenere questa informazione in maniera indiretta, attraverso il campo "titolo" presente nella sezione *profilazione sostenitore*, riferito all'intestazione utilizzata per la corrispondenza (Egregio Ingegnere, Gentile Dottoressa...). In questa sede abbiamo pensato di voler ottenere tali informazioni eventualmente tramite i Social Network per ottenere informazioni circa la professione, gli hobby e interessi personali, ma purtroppo nella maggior parte dei casi i soggetti beneficianti non risultano essere presenti sulla rete.

I campi, come detto, erano molti per cui abbiamo ritenuto opportuno in un primo momento concentrare la nostra attenzione solo su quelli che risultavano maggiormente significativi e che però erano spesso vuoti o incompleti, per cui abbiamo dovuto procedere al loro inserimento. I campi di cui parliamo sono:

- le date di nascita/decesso con i rispettivi luoghi,
- la data del testamento,
- la tipologia di testamento (che risultava incompleto perché da noi introdotto con menu a tendina per renderlo più standardizzato)
- lo stato civile (anche questo introdotto da noi con menu a tendina)
- i *coeredi*, da noi intesi come altre associazioni beneficiarie dello stesso testamento.

Si è deciso di comune accordo di tralasciare, almeno inizialmente, i dati riferiti agli aspetti economici, quali accrediti e spese di vario genere perché non è ancora tutt'oggi possibile avere un collegamento diretto tra il CRM e il software ERP utilizzato internamente per la contabilità, motivo per cui i dati inseriti in CRM sui movimenti contabilizzati non risultavano essere allineati con quelli del Controllo di Gestione.

Al di là delle problematiche sopra evidenziate la principale criticità incontrata derivava dalla presenza di dati non omogenei in quanto per molti di essi mancava una chiara tassonomia.

Così al fine di migliorare per il futuro la qualità dei dati, affinché siano meglio interpretabili sia dal personale che si occupa di analizzare questa tipologia di informazioni, che dal sistema di BI, abbiamo ritenuto opportuno definire una procedura che faccia chiarezza sia sulle responsabilità di inserimento che sulla tassonomia da utilizzare per i diversi campi informativi.

Contemporaneamente all'alimentazione del CRM, l'attenzione si è concentrata sulla progettazione del nuovo schema di datawarehouse all'interno del sistema di **Microsoft Power BI**, applicativo di Business Intelligence utilizzato per l'analisi.

Dopo aver connesso le tabelle al programma di Business Intelligence, abbiamo costruito il nuovo schema di Datawarehouse, che si presenta come uno ***schema a fiocco di neve*** per via della semplicità delle relazioni con la tabella del fatto, in quanto solo la tabella Istat Reddito è a sua volta collegata alla tabella dei Comuni Italiani.

Si presenta come uno schema poco differente rispetto a quello precedente se non per una maggiore completezza delle tabelle delle dimensioni.

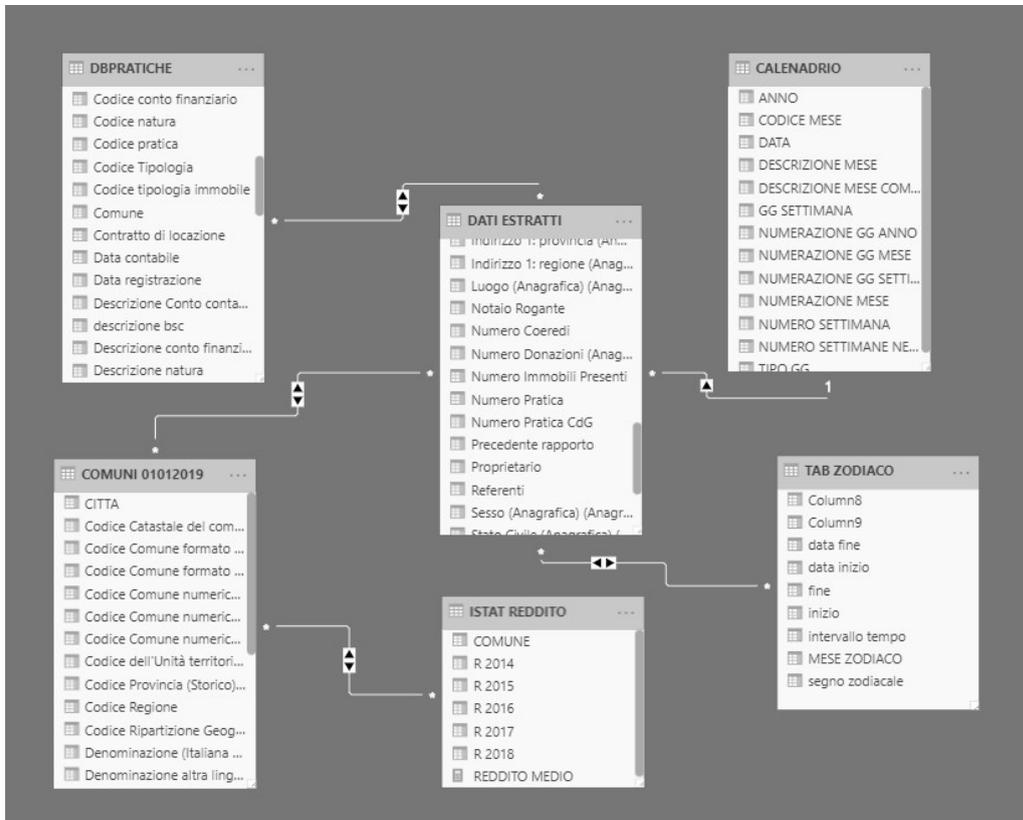


Figura 17. Lo schema definitivo di Data Warehouse creato in Microsoft Power BI.

Il programma è risultato particolarmente utile perché ha permesso di sopperire alla mancanza dei dati economici che non erano compilati in CRM e che sono stati così aggiunti al modello tramite un File Excel utilizzato dal Controllo di Gestione in cui vi sono gli accrediti e le spese riferiti ad ogni pratica. L'obiettivo è stato quello di relazionare i dati per ottenere informazioni circa il valore della singola pratica attraverso un "codice" presente in questo file e che abbiamo dovuto purtroppo riportare manualmente nel CRM per ottenere il legame.

Tornando allo scopo principale del lavoro, ossia quello di trarre informazioni dai dati per poterle utilizzare ai fini delle nuove campagne di sensibilizzazione sui Lasciti testamentari, abbiamo avviato l'analisi vera e propria.

Verificato il caricamento dei dati nel sistema di Business Intelligence, abbiamo proceduto con l'analisi costruendo le dashboard tematiche sulla base delle variabili "chiave" precedentemente selezionate.

Nella figura seguente, sono evidenziate nel riquadro in rosso una parte delle dashboard sviluppate per analizzare le principali variabili, di cui alcune molto semplici ed intuitive ma dalle quali si sono ottenuti importanti risultati che sono stati utilizzati in seguito per effettuare delle analisi incrociate tra le variabili stesse.

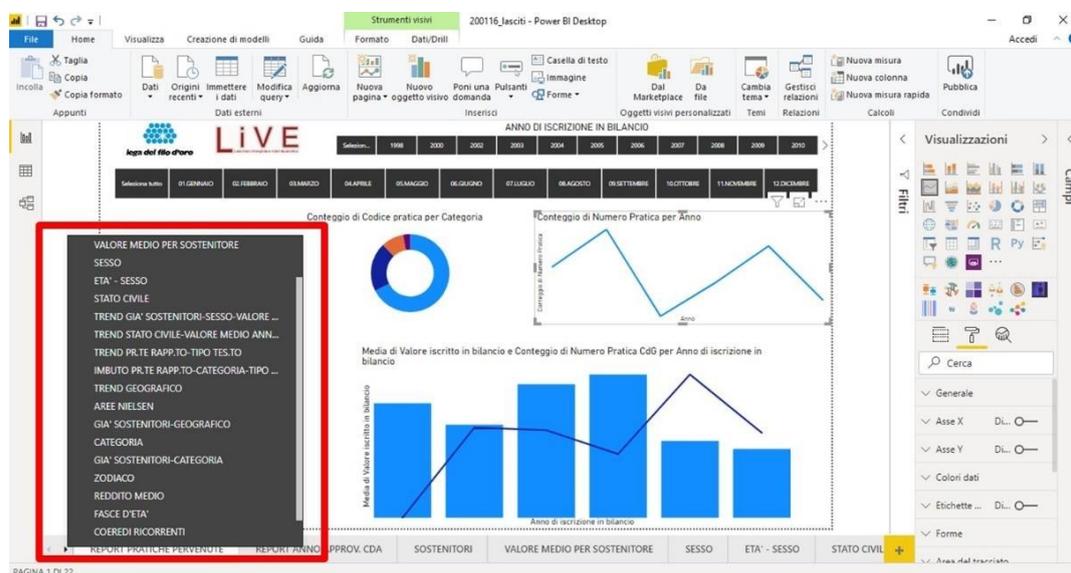


Figura 18. Nostra elaborazione.

Partendo dai dati disponibili, i principali ambiti su cui si è concentrata l'attenzione hanno riguardato lo studio delle principali variabili sociodemografiche come ad esempio il sesso, lo stato civile, l'età media in cui si fa testamento, la provenienza geografica, ottenendo talvolta risultati inaspettati.

Si è effettuato anche uno studio su variabili quantitative quali il numero di pratiche pervenute, il valore medio delle singole pratiche, il numero di immobili donati, il numero di donazioni effettuate, gli anni di fedeltà delle donazioni. Come si nota in figura però, in questa sede i risultati non vengono esplicitati in termini quantitativi (gli assi non contengono la scala numerica) per motivi di riservatezza delle informazioni e perciò vengono inseriti solo grafici a titolo di esempio.

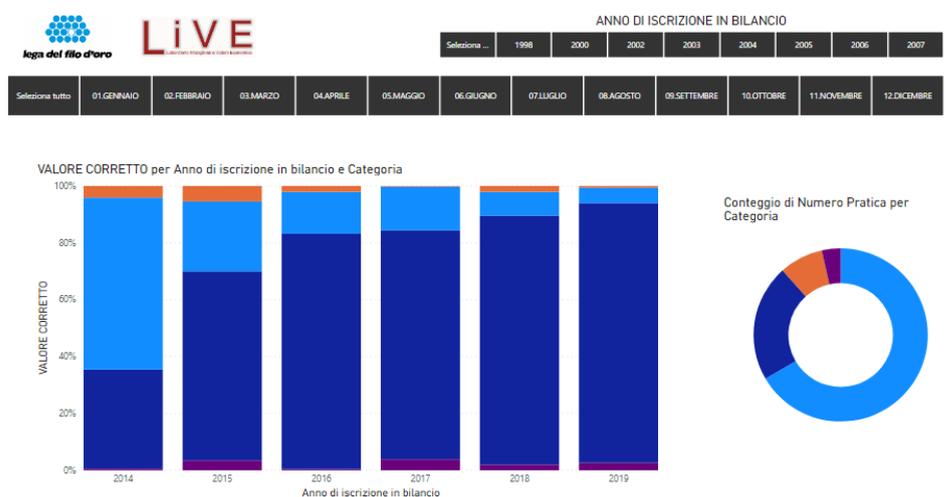


Figura 19. Il grafico a destra indica per ogni colore la categoria di lascito (donazione con atto pubblico, eredità, legato, polizza). A sinistra, la percentuale di ogni categoria sul valore totale delle pratiche per anno. (Fonte: nostra elaborazione)

Dalla figura a destra si può notare come la categoria “Azzurro” rappresenta la modalità prevalente con la quale si sottoscrive un lascito ma solo in termini assoluti, in quanto spostandosi sul grafico a sinistra si nota come in termini di valore generato è seconda dopo la categoria “Blu”. Inoltre, dal grafico a sinistra si può notare come la categoria “Blu” è cresciuta negli anni a scapito della categoria “Azzurro”, individuando così un certo trend.

Altra dashboard interessante potrebbe essere quella sulle fasce d’età.

Per costruire questa, abbiamo preso in considerazione la nuova suddivisione delle fasce d’età secondo l’Istat.

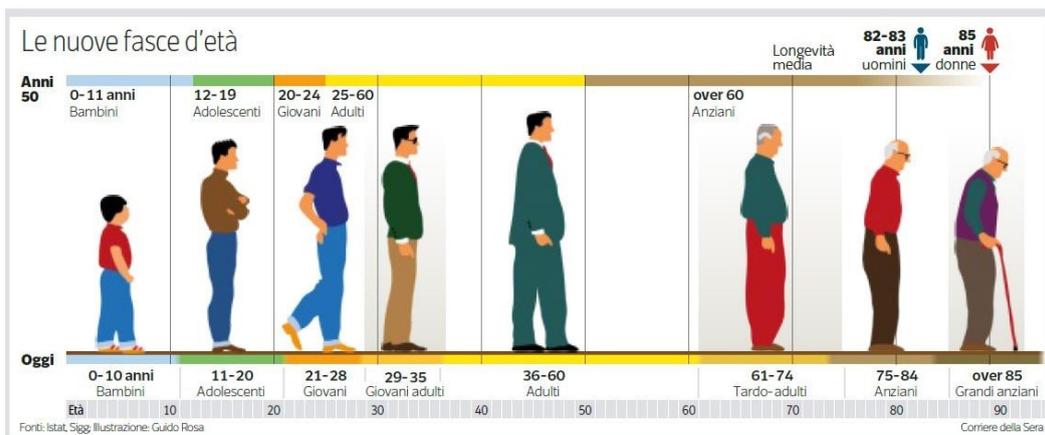


Figura 20 Nuove fasce d’età Istat (Fonte: <https://www.propensione.it/news/la-vita-si-allunga-e-con-lei-anche-la-giovinetza-27508/attachment/le-nuove-fasce-deta-corriere/>)



Figura 21. Studio delle fasce d'età prevalenti e del loro apporto in termini di valore (nostra elaborazione).

Sulla base di questa suddivisione, i lasciti sono pervenuti principalmente da due categorie di soggetti (rappresentate con il colore azzurro e blu), e in minima parte da una terza categoria. Alcune categorie hanno registrato, proprio come ci si aspettava, valori molto bassi o nulli.

Il grafico a sinistra mostra invece per ogni fascia d'età il valore della pratica rispetto al sesso e al numero di pratiche pervenute. Alcune volte il sesso viene rilevato una sola volta, ad indicare che si tratta di una simulazione non rispondente alla realtà. Il numero delle pratiche rappresentato dalla linea continua mostra come vi sia una forte differenza tra fasce d'età e ciò non sempre si rispecchia in termini di valore, rappresentato con l'altezza delle colonne.

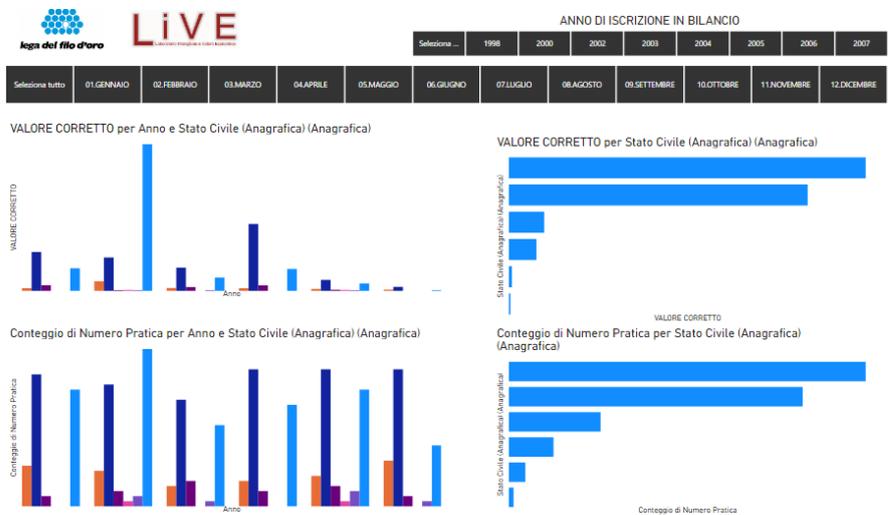


Figura 22. Analisi dello Stato Civile per numero di pratiche pervenute e in base al valore annuo registrato (nostra elaborazione).

Qui abbiamo effettuato un confronto sulle pratiche pervenute in base allo Stato Civile del soggetto beneficiante ossia, se si tratta di un soggetto Coniugato, Nubile/Celibe, Divorziato, Vedovo, Convivente, considerando che il dato in alcuni casi non è disponibile.

Si può notare nel grafico a sinistra che per ogni anno il valore delle pratiche pervenute non rispecchia quasi mai il numero di pratiche e ciò è dovuto al fatto che non sempre l'accredito del lascito viene effettuato immediatamente alla ricezione del testamento e in un'unica soluzione.

Molti altri sono gli aspetti che abbiamo considerato, qui di seguito ne citiamo alcuni.

Ad esempio, a differenza di quanto ci si attendeva, i notai non risultano rappresentare una fonte informativa fondamentale per effettuare un lascito poiché a partire dai pochi dati disponibili non si riscontra tale collegamento. La maggior parte dei testamenti viene infatti sottoscritta dal beneficiante di proprio pugno (nel 55% dei casi analizzati) facendo ricorso al testamento Olografo.

È stato interessante notare come oltre alle associazioni aderenti al Comitato Testamento solidale, all'interno del campione analizzato ci sono stati casi in cui vengono citate ulteriori associazioni, ciò ad indicare che non sempre sono i notai a promuovere tale attività ma sono gli stessi testatori che si informano volontariamente o ricevono informazioni da altri.

Abbiamo effettuato anche un'analisi di geomarketing per cercare di collegare il beneficiante alla provenienza e capire quali sono le regioni e le città più profittevoli.

È stato estremamente interessante lo studio del segno zodiacale, dato che non essendo questo un dato presente nel data warehouse e considerando che ad ogni mese non corrisponde un unico segno zodiacale, abbiamo dovuto ottenerlo in maniera indiretta. L'operazione è stata effettuata con Excel suddividendo il mese e il giorno di nascita e dando come input al sistema l'intervallo di date entro il quale rientra il segno zodiacale.

Attraverso il dato sul “precedente rapporto” abbiamo potuto riscontrare come non in tutti i casi analizzati i soggetti che hanno effettuato il lascito avevano fatto richiesta di materiale informativo o di una consulenza sull’argomento, a sottolineare come non sempre ci si informa da “fonti ufficiali”.

CONCLUSIONI

Data la centralità attribuita oggi alla conoscenza, rivestono sempre maggiore importanza i sistemi informativi aziendali, che permettono di mettere in atto una migliore “Knowledge Management”, ossia la filosofia di gestione divenuta ormai fondamentale per le imprese per ottenere vantaggi competitivi e battere i concorrenti.

Uno dei messaggi che si vuole far passare attraverso la lettura dell’elaborato è che i sistemi di Business Intelligence sono ormai divenuti fondamentali per ottenere conoscenze sui dati attualmente presenti in azienda per gestirli al meglio, ed anche disporre di una solida base di dati potenzialmente utili per il futuro, cercando di fidelizzare i clienti e stringere con essi relazioni di lungo periodo cercando di anticipare i loro bisogni.

Si parla di “gestione” perché i dati raccolti all’interno del sistema informativo vengono selezionati sulla base degli obiettivi da raggiungere ed attentamente analizzati sulla base della loro qualità. Se infatti la base dati si presenta di buona qualità permette di ottenere vantaggi in termini di efficienza nei tempi di rielaborazione dei dati e produzione delle informazioni che permetterebbe di conseguenza di anticipare le mosse dei concorrenti.

I dati estratti dai sistemi informativi permetteranno di creare il Data Warehouse necessario ad effettuare le analisi di Business Intelligence, servendosi delle tecniche di Data mining per produrre informazioni rilevanti ai fini decisionali.

L'importanza attribuita alla Raccolta Fondi ed in particolare al Lascito Testamentario da parte della Lega del Filo d'Oro ha spinto l'alta direzione a partecipare al progetto "Un ricordo tangibile" in collaborazione con "L.I.V.E Intangibles", Spin – off dell'Università Politecnica delle Marche dedicato alla Consulenza manageriale.

Lo scopo del lavoro è stato quello di individuare, tramite le dashboard create in Microsoft Power BI (di cui alcuni esempi sopra), le caratteristiche del "soggetto beneficiante" ossia colui che effettua un lascito a favore della Lega del Filo d'Oro, cercando di capire quali sono le caratteristiche più ricorrenti in questi soggetti per raggrupparli in cluster, ossia in base alle similitudini, e migliorare l'efficacia delle campagne.

Data l'introduzione abbastanza recente del sistema informativo Microsoft Dynamics 365 all'interno dell'organizzazione, in molti casi i dati già presenti non erano di elevata qualità, per cui si sono effettuate operazioni di pulizia dei dati cercando di attribuire, dove mancava, una chiara tassonomia ad alcuni campi che non si presentavano strutturati ma piuttosto descrittivi.

Per questi motivi, in accordo con il team di progetto, abbiamo stilato una procedura per l'inserimento dei dati in CRM di modo che i dati vengano inseriti in maniera piuttosto omogenea per evitare di incorrere facilmente in errori di battitura.

Attraverso le analisi sono emerse una serie di informazioni interessanti e a carattere piuttosto strategico, ottenendo spesso risultati abbastanza inaspettati, come ad esempio l'informazione sul segno zodiacale o la fascia d'età.

Nonostante le criticità incontrate, sembra comunque esserci una buona base, perlomeno sugli elementi basilari, da cui partire per poter sensibilizzare un numero sempre maggiore di sostenitori ad effettuare un gesto di estremo valore simbolico ed altamente importante per le problematiche di cui l'associazione si occupa.

BIBLIOGRAFIA

Candiotto R., *Sistemi informativi integrati*, Giuffrè Editore, 2004.

De Luca A., *Amministrazione e finanza nelle PMI innovative. Nuove metriche per valutare la performance*, Franco Angeli, Milano, 2011.

De Luca A., *CRM – Customer Relationship Management per la redditività dell'impresa. Strumenti e applicazioni*, You can print, Tricase (LE).

Inghirami I. E., *Il sistema informativo amministrativo per la governance aziendale*, Franco Angeli, 2008.

Laudon K. C., Laudon J. P., Morabito V., Pennarola F., *Management dei sistemi informativi vol. 1 Fondamenti*, Terza edizione, Pearson Prentice Hall, 2010.

Lugli G., Ziliani C., *Micromarketing, creare valore con le informazioni di cliente*, UTET, 2004.

Marchi L., Mancini D. (a cura di), *Gestione informatica dei dati aziendali*, Terza edizione, Franco Angeli, Milano, 2009.

Mauri C., *Il sistema informativo per il marketing. Dai dati alle risorse di conoscenza per la gestione delle relazioni con i clienti*, Giappichelli Editore, Torino, 2007.

Pighin M., Marzona A., *Sistemi informativi aziendali, ERP e sistemi di data analysis*, Terza edizione, Pearson, 2018.

Rezzani A., *Business Intelligence, processi, metodi, utilizzo in azienda*, Apogeo education, Maggioli editore, 2012.

Tagliavini M., Ravarini A., Sciuto D., *Sistemi per la gestione dell'informazione*, Apogeo, Milano, 2003.

Vercellis C., *Business Intelligence, modelli matematici e sistemi per le decisioni*, McGraw-Hill, 2006.

Altre fonti:

Bilancio di Sostenibilità 2018, Lega del Filo d'Oro.

Bilancio Sociale 2017, Lega del Filo d'Oro

De Carli S., *I libri della Lega del Filo d'Oro*, Vita S.p.a.

Santilli L., *La luce dentro e altri scritti su Sabina Santilli*, Cooperativa Sociale CO.A.L.A Onlus, Jesi (AN), 2003.

SITOGRAFIA

<http://www.intelligenzaartificiale.it/data-mining/>

<http://www.vita.it/it/article/2019/09/12/28-milioni-ditaliani-pensano-di-fare-un-testamento-solidale/152640/>

<https://www.bucap.it/news/approfondimenti-tematici/gestione-del-magazzino/database-data-warehouse-principali-differenze.htm>

<https://www.bucap.it/news/approfondimenti-tematici/gestione-del-magazzino/definizione-data-mart.htm>

<https://www.extrasys.it/it/smartblog/che-cosa-%C3%A8-la-business-intelligence>

<https://www.guru99.com/data-mart-tutorial.html>

<https://www.notariato.it/it/tags/testamento-solidale>

<https://www.propensione.it/news/la-vita-si-allunga-e-con-lei-anche-la-giovinezza-27508/attachment/le-nuove-fasce-deta-corriere/>

<https://www.salesforce.com/it/learning-centre/crm/what-is-crm/>

<https://www.mrwebmaster.it/sql/cosa-record 8202.html>

APPENDICE

Viene di seguito allegata la bozza di procedura dedicata all'inserimento dei dati in Microsoft Dynamics 365, il CRM utilizzato dalla Lega del Filo d'Oro per la gestione dei dati sui lasciti testamentari e le relative anagrafiche.

La bozza riprende il layout solitamente utilizzato per stilare le procedure interne.

Si tiene a precisare la non ufficialità di tale documento in quanto rappresentante un mero suggerimento per l'organizzazione, frutto della partecipazione al progetto "Un ricordo tangibile" organizzato da "L.I.V.E Intagibles".

0 INDICE

0	INDICE	1
1	SCOPO E CAMPO DI APPLICAZIONE	2
2	RESPONSABILITÀ.....	2
3	DEFINIZIONI	2
4	MODALITÀ OPERATIVE	2
4.1	Premessa	2
4.2	Creazione anagrafica.....	3
4.3	Creazione della pratica legale	4
5	MODALITÀ OPERATIVE (RAPPORTI INFORMATIVI PRECEDENTI)	6
6	STATO DELLE REVISIONI.....	6
7	FLOW CHART	7
7.1	Legato ed eredità.....	7

Emissione

Verifica
(Responsabile Sistema di Gestione Integrato)

Approvazione
(Direttore Generale)

1 SCOPO E CAMPO DI APPLICAZIONE

Lo scopo della seguente procedura è quello di fornire delle linee guida sull'utilizzo del CRM, in particolare per l'inserimento dei dati riguardanti le pratiche di Lasciti Testamentari, Donazioni con Atto Pubblico e Polizze Vita, con le relative schede anagrafiche.

La procedura si applica ogni volta che arriva un documento ufficiale (Testamento, Contratto di Polizza, Atto di Donazione) in cui viene indicata la Lega del Filo d'Oro come ente beneficiario.

2 RESPONSABILITÀ

I soggetti che si occupano dell'inserimento dati sono diversi a seconda dei casi.

In generale, i dati anagrafici sono di competenza dell'ufficio Comunicazione e Raccolta Fondi a differenza dei dati concernenti le pratiche legali che competono all'ufficio Affari generali.

L'inserimento dei dati anagrafici viene però delegato al personale degli Affari Generali se al momento di inserimento della pratica ci si rende conto che il beneficiante non è presente nell'anagrafica sostenitori.

3 DEFINIZIONI

Lascito

Definizione generica per individuare una liberalità in danaro o valori mobiliari (per esempio gioielli o polizze vita) in favore della Lega del Filo d'Oro non codificata in un testamento.

Legato

Disposizione a titolo particolare per la quale l'Ente è creditore nei confronti dell'erede/asse ereditario della somma o del bene oggetto del legato.

Eredità

Disposizione a titolo universale per la quale l'Ente subentra, totalmente o per quota, se vi sono più eredi, in tutte le situazioni giuridiche attive e passive del de cuius.

Donazione

Atto di liberalità effettuato in vita dal donante.

4 MODALITÀ OPERATIVE

4.1 Premessa

Nel momento in cui l'ufficio AG riceve notizia di una disposizione a favore della Lega del Filo d'Oro, verifica nel CRM l'eventuale esistenza della scheda anagrafica relativa al beneficiante per sapere se era già un sostenitore e/o aveva avuto un precedente rapporto¹ con l'associazione.

Accertata l'eventuale esistenza dell'anagrafica, o comunque dopo averla creata, si procede con la compilazione della scheda "pratica legale" in CRM ma solo dopo che il documento ufficiale (Testamento, Contratto di Polizza, Atto di donazione) è pervenuto all'ufficio Affari Generali, dato che all'interno di questi documenti sono contenuti i dati anagrafici del beneficiante e le

¹ Con *precedente rapporto* s'intende la richiesta di informazioni preliminari sui lasciti o l'interesse alla ricezione di materiale informativo sull'argomento, oppure una richiesta di consulenza.

informazioni sul contenuto delle disposizioni. È da sottolineare che non sempre tutte le informazioni ricercate sono disponibili, e talvolta potrebbero essere implicite nel testamento quindi dedotte solo dalla lettura di quest'ultimo.

Se il beneficiante era un Sostenitore è già presente in CRM nella sezione "ANAGRAFICA". In questa sede è importante accertarsi che il soggetto sia lo stesso verificando oltre al nome e cognome altri dati (luogo di provenienza, data di nascita...) dato che non è rara la presenza di omonimi.

Una volta individuata l'anagrafica si procede ad inserire la pratica legale proprio all'interno della scheda anagrafica cliccando in alto a destra la dicitura "pratiche legali" e cliccando poi su "aggiungi nuova entità" in alto a sinistra; da qui si apre la scheda relativa alla pratica con tutti i campi da completare.

Se il soggetto beneficiante non è presente nelle anagrafiche si tratta di un "non sostenitore", quindi si procede con la creazione di una nuova anagrafica e si seguono poi gli step successivi come sopra.

In alternativa, il CRM permette di creare prima la nuova pratica legale andando nella sezione "Ufficio Legale", quindi "Pratiche Legali", cliccando poi "Nuovo" e da qui collegare l'anagrafica riferita al beneficiante.

Si precisa che i due procedimenti sono alternativi per cui la scelta è indifferente.

La pratica legale sarà via via aggiornata con le attività svolte dai Collaboratori, i quali dovranno inserirne traccia nella sezione "Diario Pratica" presente in basso a destra della scheda pratica.

In fase di caricamento della pratica, quindi contestualmente all'arrivo del documento ufficiale, in anagrafica bisogna compilare il campo "tipo anagrafica" scegliendo dal menu apposito la voce **Lasciti** e indicare data e luogo del decesso per la disattivazione dell'anagrafica (vedi dopo).

Il beneficiante ancora in vita ha la possibilità di richiedere informazioni sui Lasciti all'apposito ufficio, il quale si occupa di curare i primi rapporti con i soggetti interessati ad effettuare un lascito. Se però questi avessero bisogno di una conoscenza più approfondita sull'argomento, quindi anche riguardante gli aspetti legali, il compito passa agli Affari Generali, che possono offrire anche un servizio di consulenza. In questo caso, qualora il soggetto interessato rilascia i propri dati anagrafici, viene inserito in CRM specificando il tipo anagrafica "**Prospect Lascito**" selezionando la voce dall'apposito elenco, ad indicare che potenzialmente potrebbe effettuare un lascito.

Altra possibilità è che il beneficiante ancora in vita (o per lui il Notaio) informa la Lega del Filo d'Oro di rientrare tra i beneficiari del Testamento, documento che inizierà a produrre i suoi effetti solo all'apertura della successione e quindi a seguito del decesso. In questo caso vale la regola di verifica dell'esistenza dell'anagrafica e quindi di creazione se inesistente, in cui bisogna indicare come "Tipo anagrafica" la voce "**Promessa lascito**". La stessa dicitura si avrà se il Testamento viene depositato presso l'ente stesso.

4.2 Creazione dell'anagrafica

Se il beneficiante non era un sostenitore e non ha avuto nessun precedente rapporto con la Lega del Filo d'Oro, l'ufficio Affari Generali procede alla creazione dell'anagrafica da collegare alla pratica legale.

In fase di creazione dell'anagrafica, i campi da compilare sono diversi.

Sono considerati obbligatori i campi: tipo account (persona fisica/società/gruppo), sesso, cognome, nome, data e comune di nascita, stato civile (ove disponibile), indirizzo di residenza, data e luogo decesso (se deceduto).

Alcuni campi vengono compilati tramite la comparsa di un menu a tendina in cui si deve selezionare la voce corretta, ma dato che non tutti i campi possono essere “automatizzati” sono comunque presenti dei campi “liberi” per i quali vengono stabilite regole di inserimento affinché i dati siano di buona qualità.

Il “Codice fiscale” va inserito nell’apposito campo dotato di controllo automatico dove una volta inserito, il dato viene verificato e il sistema segnala l’eventuale errore. Il dato viene comunque inserito manualmente per intero e dovrà essere corretto da parte dell’utente, non dal sistema.

Il campo “Titolo” fa riferimento alla forma utilizzata per le intestazioni della corrispondenza (Egregio Signore, Egregio ingegnere...). Può essere interessante per individuare la professione, essendo quest’ultimo un campo che di solito rimane vuoto perché il dato è abbastanza difficile da reperire.

Nel campo “Indirizzo completo” si scrive l’indirizzo con Via, numero civico, città, provincia e frazione e quando compare la tendina con gli indirizzi consigliati bisogna selezionare la voce corretta in modo da ottenere la normalizzazione Google.

Il luogo di nascita (e di decesso) deve essere inserito in un apposito campo legato ad una tabella dei comuni italiani con possibilità di inserire “altro” se il luogo da inserire non è presente nella lista; con l’inserimento automatico sarà preferibile avere un campo distinto per “provincia” da compilare automaticamente alla selezione del campo “luogo”. Anche per il Comune di nascita vale la stessa regola.

I campi delle date di nascita, di decesso e di redazione del testamento sono legati al calendario ma al momento dell’inserimento le date devono essere scritte per intero nel formato gg/mm/aaaa per evitare di incorrere in errori di riconoscimento dell’anno corretto da parte del programma.

Per l’inserimento dello Stato Civile vi è un apposito menu a comparsa, anche se questo tipo di informazione non sempre è disponibile. Se non si dispone dell’informazione, inserire la voce “non disponibile”.

Se il soggetto è *deceduto* (ciò vale per i testamenti e le polizze vita) bisogna spuntare il campo “deceduto” e compilare i campi di data e luogo del decesso (come sopra descritto), occorre poi disattivare l’anagrafica cliccando sui 3 puntini nella barra in alto a destra e selezionando la voce “Disattiva”, quindi “deceduto”. In questo modo l’anagrafica non viene considerata ai fini della comunicazione.

Se il *de cuius* aveva il domicilio in un luogo diverso dalla residenza si può inserire il secondo indirizzo (non obbligatorio). In questo caso si deve inserire il flag nel campo “Indirizzo di residenza diverso da spedizione” da cui compare una nuova tabella di indirizzo in cui è possibile inserire l’indirizzo del domicilio o comunque una diversa residenza, che può essere anche presso una Casa di riposo/cura. Nel caso in cui AG inserisce il secondo indirizzo è importante mantenere e non modificare il primo (salvo il caso in cui questo è errato), ossia quello di spedizione utilizzato dalla Raccolta Fondi ai fini di invio del notiziario, del materiale informativo e/o pubblicitario.

Quanto detto finora vale solo per le persone fisiche, mentre nel caso in cui il lascito venisse fatto a nome di una società i campi da compilare nell’anagrafica saranno differenti. Per la pratica legale non vi sono differenze.

4.3 Creazione della pratica legale

La creazione della scheda di pratica legale è di competenza degli Affari Generali, operazione messa in atto solo dopo aver ricevuto il documento ufficiale (testamento, atto di donazione, contratto di polizza).

Appena arriva il documento, l’ufficio AG si accerta della presenza dell’anagrafica in CRM che qualora non fosse presente viene contestualmente creata, ed entrando proprio nell’anagrafica crea il collegamento ad una Nuova pratica legale.

Per creare la scheda “pratica legale” bisogna cliccare “pratiche legali” sulla barra in alto a destra e poi “nuovo” in alto a sinistra così da visualizzare i campi da compilare.

All’interno di questa vi è una prima sezione in cui sono indicati in sintesi i dati anagrafici e cliccando proprio sul “Nome account” si apre il collegamento con la scheda anagrafica.

Appena viene creata la pratica, il sistema assegna un numero progressivo automatico non modificabile, mentre il numero pratica del Controllo di Gestione (Numero pratica CdG) ossia il codice legato al file Excel denominato “DB_Eredità” è da inserire manualmente finché non si stabilisce la regola sul codice univoco.

Anche qui ci sono dei campi che vanno compilati scegliendo la voce corretta in un set di opzioni ed è questo il caso dei campi: Categoria, Precedente Rapporto, Già sostenitore, Tipologia di Testamento.

Nel campo “data testamento” si inserisce la data seguendo le medesime indicazioni che valgono per l’anagrafica, ossia scrivendo la data nel formato gg/mm/aaaa. Qui la particolarità sta nel fatto che i testamenti potrebbero subire variazioni, essere annullati o modificati in ogni momento perciò nel caso in cui sono presenti diversi testamenti, tutti validi, si inserisce la data riferita al primo testamento in cui viene citata la Lega del Filo d’oro tra i beneficiari.

Per i Coeredi vale un discorso a parte. Attualmente i campi disponibili in CRM sono “Numero Coeredi” e “Coeredi”. Il numero dei coeredi non è quasi mai indicato, mentre nel campo “coeredi” vengono inseriti diversi dati, quali il Nome dei coeredi in quanto persone fisiche (talvolta viene aggiunto anche il numero di telefono) ma anche i nomi delle associazioni. In questo caso per evitare confusione e rendere i dati più omogenei sarebbe opportuna la suddivisione in 3 campi distinti:

- **Numero coeredi**, campo libero in cui si possono inserire solo valori numerici. Si potrebbe, almeno in prima battuta, considerare la possibilità di inserire un menu a tendina per valutare dei range per l’analisi;
- **Tipologia Coeredi**, campo con set di opzioni in cui poter scegliere tra legittimari, ascendenti, discendenti, enti;
- **Nome Coeredi**, ulteriormente diviso come campo libero per i familiari e lista per le associazioni.

Per i referenti, ossia eredi/altri soggetti che si interessano della gestione della successione oppure l’esecutore testamentario, bisogna indicare di che tipo di referente si tratta all’interno del campo libero “referenti” presente nella sezione “dettagli pratica”. Inoltre, è necessario compilare l’apposita scheda “referenti” in basso a sinistra che consiste in una semplice scheda di contatto in cui inserire informazioni circa i numeri telefonici, e-mail, indirizzo.

Per il notaio si completa una scheda di contatto simile ma che crea in automatico un’anagrafica riferita al notaio stesso (cosa che non avviene per il referente); qui l’attenzione è rivolta al tipo anagrafica da considerare per il notaio, ossia anagrafica “**Prospect Lascito**”. NB: Il notaio da inserire è quello che ha redatto il verbale di pubblicazione del testamento perciò se ci fossero documenti allegati alla pratica ma stilati da altri notai, questi ultimi non vengono presi in considerazione. Inoltre, nel caso delle Polizze manca la figura del notaio perciò il campo “Notaio rogante” resta vuoto.

Il campo “Precedente rapporto” viene compilato di default come “No”, e viene modificato come “Si” se è stato espressamente chiesto materiale informativo oppure viene richiesta una consulenza sull’argomento lasciti.

Lo stesso vale per il campo “Già sostenitore” che sarà modificato in “Si” quando il beneficiante è presente nelle anagrafiche prima ancora dell’avvento del testamento.

Il campo “responsabile interno” è compilato con il nome del Direttore Affari Generali che rappresenta l’unica scelta possibile tra le opzioni, mentre la gestione della pratica viene assegnata in un primo momento allo stesso Direttore, il quale poi provvederà ad assegnarla

ad uno dei suoi collaboratori ma solo successivamente alla delibera di accettazione da parte del CDA.

Le date di creazione pratica e anagrafica, di chiusura della pratica, riapertura e chiusura definitiva sono compilate automaticamente nel momento in cui viene effettuata la relativa operazione.

Una volta inseriti questi dati, nei limiti della loro disponibilità, si salva la pratica e si inserisce all'interno del diario pratica la data di ricezione del testamento con relativa descrizione dell'evento.

Altra informazione da inserire in fase di creazione della pratica fa riferimento al campo libero "Stato di avanzamento pratica" che in un primo momento sarà "In preparazione per CDA del gg/mm/aaaa" e che verrà modificato di volta in volta con le vicende relative alla pratica stessa.

Dopo aver effettuato i diversi adempimenti riguardanti il lascito ed incamerato quanto previsto, nel caso in cui non si prevede di ottenere altro, la pratica viene chiusa al livello 1 con il tasto "**Chiudi pratica**" e dovrà essere disattivata cliccando sul tasto "**Disattiva**".

Una volta disattivata, la pratica **NON** potrà essere riattivata in un secondo momento salvo casi eccezionali.

Qualora sussista l'aspetto patrimoniale, lo stato della pratica si chiude al livello 1 e si apre al livello 2 in stato di "extra gestione" finché rimane l'immobile iscritto a patrimonio.

Una volta venduto l'immobile la pratica verrà chiusa anche al livello 2 e si può procedere alla disattivazione definitiva.

5 MODALITÀ OPERATIVE (RAPPORTI INFORMATIVI PRECEDENTI)

Se la persona interessata ad effettuare un Lascito contatta la Lega per richiedere informazioni sull'argomento Lasciti, è compito della Raccolta Fondi - divisione Lasciti di intrattenere i primi rapporti.

Quando l'interessato è un sostenitore, va inserita traccia nel CRM indicando la data e il tipo di contatto.

Nel caso in cui la richiesta di informazioni o l'invio di materiale perviene da un soggetto che non è sostenitore, quindi non presente nelle Anagrafiche, si crea in CRM una nuova anagrafica specificando come tipo anagrafica **Prospect Lascito**. Ciò però non sempre è possibile perché alcune volte non si ha la disponibilità neppure dei dati fondamentali; converrebbe quindi creare l'anagrafica quando il rapporto diventa più formale.

Potrebbe accadere che lo stesso beneficiante che informa la Lega di essere tra i beneficiari del proprio testamento. In questo caso si tratta comunque di precedente rapporto, quindi si inserisce l'anagrafica (se non esiste già) e si specifica il tipo anagrafica **Promessa Lascito**.

6 STATO DELLE REVISIONI

Rev	Data	Motivo della Revisione
0		
		-
		-
		-

7 FLOW CHART

7.1 Legato ed eredità

