



# **UNIVERSITÀ DEGLI STUDI DI PADOVA**

---

**Dipartimento di Tecnica e Gestione dei Sistemi  
Industriali**

**Corso di laurea Triennale in Ingegneria Gestionale**

## **BUSINESS INTELLIGENCE PER LE PMI: STATO DELL'ARTE E COMPARAZIONE SOFTWARE**

RELATORE: CH.MO PROF. BONOLLO GIULIANO

LAUREANDO: TERZO FRANCESCO

---

ANNO ACCADEMICO 2015-2016



# Indice

Indice.....	3
Introduzione.....	5
<b>Capitolo 1. Business Intelligence e Data Warehouse.....</b>	<b>7</b>
1.1. Componenti della BI.....	9
1.1.1. ETL e Data Warehouse (back end).....	11
1.1.2. Funzionalità (front end).....	13
1.1.2.1. <i>Reporting</i> .....	14
1.1.2.2. <i>Analisi Dinamiche (OLAP)</i> .....	14
1.1.2.3. <i>Simulazione (What if)</i> .....	16
1.1.2.4. <i>Data Mining</i> .....	16
1.2. I Big Data.....	18
1.2.1. Le PMI e i Big Data.....	20
<b>Capitolo 2. Panoramica del mercato.....</b>	<b>23</b>
2.1. Prodotti Storici.....	23
2.1.1. Oracle Corporation.....	24
2.1.2. Microsoft Corporation.....	27
2.1.3. IBM.....	29
2.1.4. SAP.....	33
2.1.5. MicroStrategy.....	35
2.2. Prodotti Indipendenti.....	38
2.2.1. Qlik.....	39
2.2.2. TIBCO Spotfire.....	41
2.2.3. Tableau.....	43

2.3. Prodotti Open Source.....	46
2.3.1. Pentaho.....	47
2.3.2. TIBCO Jaspersoft BI.....	50
2.3.3. Eclipse BIRT Project.....	51
<b>Capitolo 3. Comparazione software.....</b>	<b>55</b>
3.1. PMI: Definizioni.....	55
3.2. Tecnologie database.....	56
3.3. Parametri Software Selection.....	62
3.3.1. Parametri.....	62
3.3.2. Punteggi e valutazioni.....	66
3.4. Confronto.....	67
3.4.1. Prodotti Storici.....	71
3.4.2. Prodotti Indipendenti.....	77
3.4.3. Prodotti Open Source.....	81
3.5. Alcune soluzioni e suggerimenti.....	84
<b>Capitolo 4. Conclusioni.....</b>	<b>89</b>
Bibliografia.....	91

## ***Introduzione***

La Business Intelligence, concetto che verrà approfondito nel primo capitolo, è sempre più una realtà che le aziende si trovano a dover affrontare per rimanere competitive nel loro settore. Inoltre, mentre anni fa era un fenomeno che interessava solo le grandi realtà industriali, oggi interessa anche le medie e le piccole imprese. Proprio queste aziende, a causa delle poche risorse che spesso sono in grado di dedicare a quest'area, devono effettuare uno studio e un'analisi molto approfonditi delle esigenze, in prima analisi, e dei prodotti offerti dal mercato dei software di BI, in seconda analisi. Questa tesi ha lo scopo di approfondire entrambi gli aspetti, offrendo sia una panoramica della Business Intelligence e di tutte le sue aree, sia un'analisi del mercato che risulta essere molto complesso e affollato. Per questo motivo verranno analizzate tre diverse categorie di prodotti: i prodotti storici, i prodotti indipendenti e i prodotti open source. La suddivisione è del tutto arbitraria e stabilita per questo lavoro, ma sono possibili moltissime altre suddivisioni, come ad esempio quella offerta dall'analisi di Gartner che in base a due criteri divide i prodotti in quattro categorie che verranno approfondite nel terzo capitolo: leaders, challengers, visionaries, niche players.

In questa tesi, i criteri con cui sono stati suddivisi i prodotti sono:

- Storia e diffusione nel mercato
- Fatturato
- Licenze

I prodotti storici sono quelli considerati leader del mercato, caratterizzati da enormi fatturati, costi elevati per il cliente e con una diffusione molto estesa e consolidata. I prodotti indipendenti sono quelli specializzati in un determinato settore e non sono parte dell'offerta di uno dei prodotti storici, ma hanno avuto la capacità di diffondersi autonomamente. Gli open source, infine, sono quelli che all'interno della propria offerta presentano una soluzione gratuita e che offrono anche licenze aperte e non solo commerciali.

Al termine della panoramica del mercato è necessario effettuare un confronto tra i vari prodotti per vedere, una volta stabilite le esigenze dell'azienda, quali rispettano meglio le caratteristiche del business. Risulta sostanzialmente impossibile determinare una soluzione che vada bene per una determinata tipologia di aziende, in quanto anche realtà simili hanno esigenze e logiche aziendali che possono essere completamente diverse. Quello che verrà fatto nel terzo capitolo sarà presentare alcune soluzioni, tra le

moltissime possibili, che potrebbero adattarsi alle esigenze di una piccola-media impresa (PMI).

Quello che è importante sottolineare è che ogni azienda deve cercare di individuare la soluzione che meglio si adatta alla situazione dell'impresa, alle logiche adottate e alle condizioni del mercato attraverso tutto il processo di studio e analisi. Solo in questo modo le PMI possono guardare al futuro con possibilità di consolidamento e crescita.

# CAPITOLO 1

## Business Intelligence e Data Warehouse

Ogni giorno una piccola media impresa deve gestire un volume di dati in costante aumento. Di conseguenza cresce sempre di più la necessità di raccogliere, gestire e analizzare tutte le informazioni di cui dispone. Un concetto ormai indispensabile per ogni tipologia di azienda è il sistema informativo, inteso come un insieme di strumenti, risorse e procedure che consentono la gestione delle informazioni. Parte fondamentale all'interno di questo meccanismo è svolta dal sistema informatico, cioè l'insieme dei sistemi hardware e software presenti in azienda. Grazie ad una infrastruttura ben costruita, l'azienda è in grado di raccogliere dati sia da fonti interne, come i vari reparti e le informazioni storiche, sia da fonti esterne per integrarle in un unico sistema centralizzato. Tutti questi dati raccolti e integrati possono successivamente essere analizzati con fini di controllo, simulazione, previsione e molto altro. Un sistema informatico ben strutturato e all'avanguardia può fare la differenza soprattutto per le piccole e medie imprese che si trovano in un'area del mercato molto difficile. Per tutti questi motivi si può intuire l'importanza di una scelta corretta dei vari software da utilizzare per poter trarre tutti i benefici derivanti dall'organizzazione dei dati al fine di ricavarne delle informazioni aggiuntive. Oltre all'implementazione iniziale di questi sistemi, sono necessarie anche la manutenzione e l'aggiornamento costante per rimanere al passo con le tecnologie offerte. Come definito dall'Information Technology Association of America "lo studio, progettazione, sviluppo, applicazione, implementazione, sostegno e gestione di sistemi informativi basati sui computer" prende il nome di *Information Technology* (IT), intesa come l'utilizzo di elaboratori per la memorizzazione, recupero, trasmissione e manipolazione dei dati.

Il termine Business Intelligence (BI) viene coniato per la prima volta nel 1958 quando Hans Peter Luhn, ricercatore tedesco che al tempo lavorava presso IBM, elabora un sistema automatico in grado di diffondere informazioni tra le varie aree di una stessa organizzazione, sia essa industriale, scientifica o governativa.

Il concetto di BI è spesso utilizzato in modo improprio, o comunque impreciso, anche se oggi può essere associato ad un insieme di procedure, tecnologie e processi atti a raccogliere, integrare e analizzare i dati provenienti dalle varie aree di un'azienda al fine di trasformarli in informazioni utilizzabili per le decisioni strategiche di business. La

trasformazione dei dati in informazioni è una delle esigenze delle aziende che ha portato alla diffusione di questi sistemi, assieme alla necessità di prendere decisioni in tempi molto stretti utilizzando un approccio razionale.

Per tutti questi motivi risulta utile evidenziare la differenza tra i concetti di dato e informazione. I dati che vengono acquisiti dalle varie aree aziendali sono analitici, cioè ad elevato livello di dettaglio, raccolti in tempo reale e aventi orizzonte temporale di breve termine (2-3 mesi). Le informazioni invece sono sintetiche, orientate all'analisi rispetto diverse dimensioni (contengono anche la dimensione temporale) e con un orizzonte di lungo termine (fino a 10 anni). Le informazioni si ottengono nel momento in cui più dati elementari vengono messi in relazione tra loro; in questo modo da grandi volumi di dati non organizzati si possono ottenere sottoinsiemi pertinenti alle figure incaricate a prendere le decisioni di business (*decision maker*), in una forma in cui queste persone possono facilmente metterli in relazione. Questi manager generalmente prendono decisioni in base alle proprie esperienze, alla percezione dell'attività aziendale e alle informazioni disponibili. Spesso esperienze, percezioni e strategie impiegate nel processo decisionale sono alquanto statiche: cioè cambiano in modo molto lento. Le informazioni, al contrario, sono sempre nuove, il che significa che cambiano molto frequentemente e in modo veloce. Risulta di conseguenza difficile avere il controllo su questi cambiamenti e mantenerne il passo.

La BI aiuta un migliore processo decisionale analizzando se, di fatto, le azioni intraprese stanno andando di pari passo con gli obiettivi dell'azienda. Decidere qual è una scelta migliore per un'organizzazione si consegue meglio con un insieme di obiettivi chiaramente stabiliti e un piano per conseguirli. La BI non riceve semplicemente il piano e lo utilizza come bilancia per misurare la qualità delle decisioni, ma ha un ruolo principale nella creazione di queste strategie e di questi piani. Si tratta di prendere più velocemente migliori decisioni, e le decisioni più strategiche sono quelle in cui la BI è più indispensabile.

Un possibile rischio della BI è di creare all'interno di un'organizzazione aziendale una sorta di "anarchia informativa" che consiste nell'utilizzo incontrollato dei dati da parte di chiunque. Risulta evidente come questo possa essere problematico nel processo di coordinamento di un'impresa al fine di raggiungere obiettivi comuni. Di conseguenza, una soluzione di BI può essere sfruttata nel migliore dei modi quando l'informazione distribuita è anche fortemente standardizzata, il che significa non una limitata libertà di accesso ai dati, ma la massima disponibilità di informazioni certificate. Tutta l'infrastruttura informativa che sta alla base della BI deve quindi essere tale da

rappresentare l'unica fonte di informazioni in grado di reperire dati tempestivamente e facilmente, con un'elevata profondità storica che permette un'analisi a diversi livelli di aggregazione e secondo molteplici dimensioni.

## 1.1. Componenti della BI

Una soluzione di Business Intelligence implica una tecnologia in grado di soddisfare le esigenze degli utenti e in particolar modo dei decision maker.

Fino a qualche anno fa, la funzione svolta in un'azienda dalle basi di dati (*database*) era quella di memorizzare i dati operazionali, cioè dati provenienti generalmente da processi gestionali. Per poter estrarre delle informazioni è necessario, quindi, che l'amministratore di questo database esegua delle interrogazioni mirate allo scopo dell'analisi. Il problema si presenta nel momento in cui è necessario coordinare le informazioni, sempre più numerose, provenienti da più database diversi, magari presenti in aziende diverse. Verso la fine degli anni '80, inizio anni '90, il ruolo delle basi dati ha cominciato a cambiare, mutando da semplice raccogliitore di dati (*repository*) a strumento attraverso il quale, grazie all'utilizzo di tecniche informatiche, è possibile estrapolare informazioni di supporto alle decisioni aziendali. Si è arrivati al concetto di processo di *data warehousing*, alla base del quale l'idea è quella di separare l'elaborazione di tipo analitico da quella legata alle transazioni, costruendo un nuovo raccogliitore di informazioni che integri i dati elementari provenienti da sorgenti di natura differente e, una volta organizzati, li renda disponibili a vari livelli di aggregazione.

Una definizione dei data warehousing la danno Golfarelli e Rizzi nel libro "Data Warehouse, teoria e pratica della progettazione" (2006):

*"Il data warehousing è una collezione di metodi, tecnologie e strumenti di ausilio al cosiddetto "lavoratore della conoscenza" (knowledge worker: dirigente, amministratore, gestore, analista) per condurre analisi dei dati finalizzate all'attuazione di processi decisionali e al miglioramento del patrimonio informativo".*

Al centro di questo processo, il *data warehouse* (DW), costituisce la base dati su cui si appoggia tutta l'analisi dei dati in esso raccolti. Una definizione ci viene data da William H. Inmon, informatico statunitense fondatore della Corporate Information Factory, il quale fu il primo a parlare esplicitamente di DW. Inmon definisce il data warehouse (letteralmente, magazzino di dati) come una collezione di dati di supporto per il processo decisionale che presenta le seguenti caratteristiche:

- *Integrata*: nel DW confluiscono dati provenienti da più sistemi transazionali o addirittura da sistemi informativi esterni all'azienda. L'obiettivo è quello di restituire una visione unificata e coerente di tutti questi dati;
- *Orientata al soggetto*: il DW è orientato a temi aziendali specifici (acquisti, vendite, ...). L'obiettivo non è più quello di eliminare la ridondanza mediante la normalizzazione, ma quello di fornire informazioni organizzate in modo tale da favorire la comprensione e l'analisi.
- *Variabile nel tempo*: a differenza dei dati operazionali che hanno un orizzonte temporale piuttosto limitato, il DW deve permettere analisi che spazino su un intervallo temporale di alcuni anni.
- *Non volatile*: è una caratteristica che assicura l'integrità del dato. L'aggiornamento del DW viene generalmente eseguito mentre il DW è fuori linea il che fa sì che possa essere considerato come un database di sola lettura.

Ulteriore caratteristica del DW, che lo differenzia dai database operazionali, è la modalità di interrogazione. Nei DB operazionali, le interrogazioni eseguono transazioni che in genere leggono e scrivono un numero ridotto di record. Questo tipo di elaborazione viene definita *On-Line Transactional Processing (OLTP)*. Al contrario, il tipo di elaborazione che caratterizza i DW viene detto *On-Line Analytical Processing (OLAP)*, basato su un'analisi dinamica e multidimensionale di una quantità elevata di record.

Per capire meglio il funzionamento dei vari tool di BI è necessario comprendere e analizzare la struttura di un DW. Principalmente esso si articola in quattro livelli attraverso i quali avviene il flusso dei dati dai sistemi sorgenti fino alla struttura finale. I quattro livelli sono:

- *Livello delle sorgenti*: come già detto, i dati elementari possono provenire da fonti dati differenti come database operazionali o altre tecnologie.
- *Livello di alimentazione*: i dati devono essere ripuliti, organizzati e resi omogenei, in modo da poter essere confrontati tra di loro. Gli strumenti ETL (*Extraction, Transformation and Loading*) permettono di estrarre i dati dalle sorgenti, ripulirli, trasformarli e in seguito caricarli nel DW.
- *Livello del warehouse*: tutte le informazioni vengono raccolte in un unico contenitore che altro non è che un database dedicato. Accanto al DW è presente un ulteriore contenitore all'interno del quale vengono inseriti i *metadati (data about data)*, cioè informazioni aggiuntive sui dati contenuti nel DW.

- *Livello di analisi*: permette la consultazione rapida ed efficace dei dati integrati con lo scopo di realizzare report, eseguire analisi e svolgere simulazioni.

Successivamente i vari livelli verranno trattati in maniera più approfondita per permettere una miglior comprensione dell'architettura informatica che sta alla base di un sistema BI.

### 1.1.1. ETL e Data Warehouse (back end)

Il livello di alimentazione è probabilmente il più importante e il più delicato in quanto, affinché tutta la parte di analisi e reportistica funzioni e sia efficiente, è necessario che i dati su cui si basa il DW siano puliti e coerenti sia tra loro che con le direttive aziendali e le regole di business. Spesso viene introdotto un ulteriore livello, detto *livello dei dati riconciliati*, che materializza in una posizione di stoccaggio temporanea i dati estratti dalle sorgenti in modo da poterli trasformare e riconciliare prima di caricarli nel DW. Questa zona per la trasformazione dei dati viene definita *staging area* e al suo interno avvengono tutti i processi ETL (*Extraction, Transformation and Loading*). Tra le diverse fasi del processo di data warehousing risulta essere tra quelle più complesse e tecnicamente impegnative. I tool ETL svolgono principalmente quattro funzioni:

- *Estrazione*: in questa fase i dati rilevanti ai fini del DW vengono estratti dalle sorgenti. La scelta dei dati di interesse da estrarre dipende principalmente dalla loro qualità, determinata dalla presenza e dall'accuratezza dei vincoli a livello delle sorgenti e dal contesto in cui sono posizionati.
- *Pulitura*: lo scopo di questa fase è quella di migliorare la qualità dei dati, normalmente piuttosto scadente a livello delle sorgenti. Le principali cause di "sporcizia" dei dati sono la presenza di duplicati, dati mancanti, inconsistenti ed errati. La pulizia dei dati consiste nella correzione degli errori in base alla *naming convention*, cioè regole grazie alle quali si decide di utilizzare una nomenclatura uniforme per tutto il DW.
- *Trasformazione*: è la fase centrale del processo di riconciliazione e consiste nel convertire i dati dal formato che avevano in sorgente al formato del DW. Negli strumenti ETL la fase di pulizia e di trasformazione sono generalmente sovrapposte.
- *Caricamento*: consiste nella propagazione degli aggiornamenti al DW. Il caricamento dei dati può avvenire per *refresh* (caricamento totale) o per *update* (caricamento delle sole modifiche).

Tutte queste sono funzionalità che ci si aspetta di trovare in uno strumento di back end di BI. Altro aspetto importante nella fase di ETL e quindi nei rispettivi tool, è la presenza di un repository di metadati, cioè dati utilizzati per descrivere altri dati. Essi indicano la provenienza, il valore, l'utilizzo e le funzioni dei dati memorizzati nel DW.

Una delle parole chiave del livello del warehouse, ma più in generale di tutto il data warehousing, è multidimensionalità. Grazie alla sua semplicità e intuitività, il modello multidimensionale è sempre più utilizzato per la rappresentazione logica dei dati nei DW. Questo modello concettuale si basa sui *fatti* del mondo aziendale, cioè oggetti di interesse per il processo decisionale (vendite, acquisti, produzione...) e le occorrenze di un singolo fatto corrispondono a *eventi* accaduti. Per ciascun fatto si possono individuare *misure*, cioè proprietà numeriche che ne descrivono un aspetto quantitativo di interesse, e *dimensioni* che corrispondono a proprietà con dominio finito e che ne rappresentano una coordinata per l'analisi. È proprio il concetto di dimensione che ha dato origine alla metafora, ormai molto diffusa, del cubo per la rappresentazione dei dati multidimensionali. Secondo questa metafora gli spigoli del cubo rappresentano le dimensioni di analisi (se le dimensioni sono maggiori di tre si parla di *ipercubo*) e le celle corrispondono agli eventi caratterizzati da un valore per ogni misura. Ogni cubo è incentrato su un fatto di interesse per il processo decisionale del quale si può eseguire un'analisi rispetto ad una dimensione a diversi livelli di dettaglio. Questo modello logico, insieme al modello relazionale, rappresenta l'approccio principale all'implementazione di un DW.

Gli strumenti di reportistica e di analisi sono strettamente collegati al livello del warehouse dal quale possono ricevere i dati dal DW stesso (che prende il nome di DW primario) oppure dai cosiddetti *data mart*, cioè sottoinsiemi del DW primario, del quale ne conservano la struttura, corrispondenti ad una particolare divisione aziendale o una particolare area di business. I data mart si dividono in due categorie: indipendenti (nel caso in cui non vi sia alcuna tabella condivisa tra più data mart ma siano quindi duplicate) e federati (aventi tabelle condivise). La suddivisione del DW in sottoinsiemi rappresentanti i diversi fatti di interesse, è un modo per sviluppare un DW gradualmente implementando un data mart alla volta senza correre il rischio che tutto il sistema diventi obsoleto prima ancora di averlo implementato completamente e messo in funzione.

### 1.1.2. Funzionalità (front end)

Le informazioni immagazzinate e ordinate all'interno del DW, e quindi dei data mart, sono rese disponibili agli utenti che possono effettuare interrogazioni e analisi attraverso quelli che vengono chiamati strumenti di front end, quali report generator, strumenti OLAP, Data Mining e strumenti per l'analisi What if.

La business intelligence aiuta i manager a prendere migliori decisioni più velocemente grazie alle analisi che possono effettuare sul DW. In generale, è possibile considerare l'analisi come un ciclo articolato in quattro fasi:

- *Analisi*: consiste nell'individuare e nel comprendere la problematica da affrontare. In questa fase il decision maker elabora un modello del fenomeno da analizzare individuando le aree d'interesse e i fatti coinvolti.
- *Comprensione*: in questa fase le informazioni estratte dall'analisi vengono utilizzate per raggiungere una conoscenza più approfondita del problema da trattare.
- *Decisione*: le conoscenze acquisite negli step precedenti vengono tradotte in decisioni strategiche e di business che successivamente verranno sviluppate in azioni pratiche da mettere in atto.
- *Misura*: si esegue una misurazione delle prestazioni e dei cambiamenti apportati dalle azioni intraprese nella fase precedente. E' una fase molto importante che permette di capire i benefici derivati dalle decisioni prese e individuare eventuali interventi correttivi al fine di migliorare ulteriormente l'intero processo.

Nel campo dell'informatica il front end consiste nella parte di un sistema software che gestisce l'interazione con l'utente. Gli strumenti front end nel contesto della BI, di conseguenza, sono tool che, attraverso interfacce grafiche, permettono agli utenti di interagire con il DW e quindi di accedere ai dati contenuti al suo interno. Le principali tipologie di analisi, e quindi di funzionalità di questi strumenti, sono:

- *Reporting*
- *Analisi dinamiche (OLAP)*
- *Simulazione (What-if)*
- *Data Mining*

### 1.1.2.1. Reporting

Questa tipologia di analisi è orientata agli utenti che hanno la necessità di accedere a informazioni strutturate in maniera autonoma, in modo da poter creare da soli le interrogazioni e le analisi. Le informazioni che ne derivano vengono successivamente organizzate in presentazioni grafiche e talvolta dinamiche, definite *report* (rapporti).

Un *report* è definito da un'interrogazione e da una presentazione. L'interrogazione consiste nella selezione e nell'estrazione dal database di una serie di dati di interesse per il determinato scopo del report stesso. La presentazione può essere in forma tabellare oppure grafica (istogrammi, diagrammi a torta, ...). Oltre alla ricchezza delle modalità di presentazione dei dati, altri due aspetti fondamentali nella valutazione di uno strumento di reporting sono la flessibilità dei meccanismi per la loro distribuzione e la presenza di uno strato semantico logico che permetta di nascondere la complessità della struttura fisica del database. In questo modo l'utente non è costretto ad interfacciarsi con l'architettura fisica del DB, ma con una sua riorganizzazione logica semplificata.

Si possono individuare principalmente tre tipologie di report:

- *Report statici*: il potenziale informativo in termini di misure e dimensioni di analisi, così come la struttura e il lay-out, sono definiti a priori e non modificabili dall'utente.
- *Report dinamici*: è possibile un'analisi dinamica rispetto alle dimensioni di interesse. Si dividono a loro volta in due categorie: report desk, modificabili dall'utente solamente in locale e mediante un software appropriato, e report web, creati e salvati attraverso un'applicazione web.
- *Dashboard (cruscotto)*: sottoinsieme di reporting che presenta dati e informazioni relative alle prestazioni di un'organizzazione o di una specifica funzione, in modo grafico, sintetico e con un certo grado di interattività. I cruscotti sono molto utilizzati per la presentazione in tempo reale dei dati relativi ad una particolare area d'interesse, ma sono leggermente carenti dal punto di vista del reporting.

### 1.1.2.2. Analisi Dinamiche (OLAP)

Con la sigla OLAP (*On Line Analytical Processing*) si intende la possibilità messa a disposizione dell'utente di costruire, modificare e personalizzare analisi. Mentre gli utenti degli strumenti di reportistica svolgono un ruolo sostanzialmente passivo, in quanto viene loro fornita una serie di strutture predefinite per i report, gli utenti OLAP hanno la

possibilità di gestirsi una sessione di analisi e di esplorazione dei dati. La poca esperienza che generalmente gli utenti hanno a disposizione, la complessità della struttura dei dati e la profonda conoscenza necessaria di quest'ultimi, rendono fondamentale il ruolo dello strumento utilizzato che deve essere di conseguenza intuitivo e flessibile.

Le principali operazioni che gli strumenti OLAP sono in grado di svolgere durante una sessione di analisi attraverso le informazioni contenute in un cubo, sono:

- *Roll-up* (letteralmente, rotolare): operazione che permette di aumentare l'aggregazione dei dati eliminando un livello di dettaglio da una gerarchia.
- *Drill-down* (letteralmente, trivellare): consiste nella funzione opposta al roll-up in quanto diminuisce l'aggregazione dei dati aggiungendo un livello di dettaglio.
- *Slice-and-dice* (letteralmente, tagliare a fette e a cubetti): consiste nella possibilità di esprimere una variabile rispetto a qualsiasi incrocio di dimensioni ad essa associate.
- *Drill-across*: possibilità di stabilire un collegamento tra due o più cubi al fine di eseguire una comparazione dei dati.

I termini ROLAP (*Relational OLAP*) e MOLAP (*Multidimensional OLAP*) rappresentano i due approcci principali per l'implementazione dei DW che si basano rispettivamente su una rappresentazione relazionale e una multidimensionale dei dati. Il vantaggio principale nell'utilizzare un modello relazionale sta nella diffusione molto vasta dei RDBMS (*Relational Database Management System*, cioè sistemi per la gestione di basi di dati relazionali). D'altro canto, ne risentono le prestazioni a causa delle complesse operazioni di join che è necessario eseguire sulle tabelle del database. Viceversa, un sistema MOLAP si basa su un modello logico nel quale i dati e le operazioni possono essere rappresentati direttamente. Il principale vantaggio sta nell'esecuzione delle operazioni in modo semplice e naturale, senza la necessità di complesse operazioni di join e, di conseguenza, in prestazioni nettamente migliori rispetto ai sistemi ROLAP.

Oltre all'analisi dei dati, altro concetto fondamentale per la BI è il *data discovery*, cioè una serie di soluzioni basate su tecniche di analisi più potenti in grado di confrontare dati provenienti anche da sorgenti eterogenee, al fine di prevedere eventi futuri e individuare informazioni nascoste. In seguito verranno analizzate due funzionalità front-end molto importanti all'interno della BI.

### 1.1.2.3. Simulazione (*What if*)

L'analisi *What-If* è una tecnica di simulazione che, dato un modello matematico di un determinato sistema reale, permette di eseguire una previsione del suo comportamento futuro sulla base dei dati contenuti nel DW. Questo tipo di analisi permette di superare il limite delle interrogazioni OLAP (che consideravano esclusivamente il passato), per volgere l'attenzione agli eventi futuri. Il primo passo per ottenere delle simulazioni e delle previsioni attendibili e quanto più precise, consiste nel costruire un modello in grado di rappresentare nel migliori dei modi la realtà. Chiaramente quanto più il modello è dettagliato e preciso, tanto più l'analisi avrà successo.

### 1.1.2.4. Data Mining

Le tecniche di *Data Mining* rappresentano un ulteriore passo conoscitivo nell'analisi dei dati in quanto consentono di individuare informazioni "nascoste" nei dati. Più precisamente, è possibile definire il Data Mining come un processo analitico che permette di esplorare grosse quantità di dati con lo scopo di individuare e rendere visibili le associazioni nascoste tra le informazioni, attraverso la ricerca di regolarità consistenti e relazioni sistematiche (*pattern*) tra variabili.

Il *data discovery* è un processo iterativo schematizzabile attraverso quattro fasi:

1. *Selezione e campionamento*: viene individuato il fatto di interesse e, di conseguenza, tutti i dati collegati al cubo corrispondente, assieme ad informazioni che possono provenire da altre sorgenti.
2. *Trasformazione e pulitura*: i dati vengono ripuliti e resi coerenti.
3. *Data Mining*: viene lanciato l'algoritmo grazie al quale sarà possibile individuare i pattern.
4. *Valutazione*: controllo e validazione dei pattern individuati

Questo tipo di analisi è molto complessa e richiede conoscenze in ambito matematico e statistico maggiori rispetto alle altre tecniche front end. Il valore aggiunto che il Data Mining può dare, però, è davvero importante e spesso determinante per un'azienda per differenziarsi dai diretti concorrenti nel mercato.

Il Data Mining comprende il concetto, sempre più diffuso in ambito aziendale, di analisi predittiva (*predictive analytics*), intesa come strumento di previsione degli eventi futuri. La predictive analytics è in grado di migliorare la comprensione del business,

contribuendo a prevedere il comportamento degli utenti e le performance dell'organizzazione. Un modello predittivo identifica le relazioni all'interno dei dati per prevedere ad esempio le tendenze di mercato e i comportamenti dei clienti, offrendo dei vantaggi notevoli in ambito di strategie di mercato e ottimizzazioni aziendali. Successivamente, grazie alle informazioni ricavate da questo tipo di analisi, diventa più semplice prendere delle decisioni di business. L'analisi predittiva risulta quindi essere il punto di partenza dell'analisi prescrittiva che consiste nella definizione delle azioni concrete da intraprendere. Le tecniche utilizzate per condurre le analisi predittive possono essere raggruppate in due tipologie:

- *Tecniche di regressione*: consistono nella creazione di un'equazione matematica come modello per rappresentare le interazioni tra i diversi dati che si stanno analizzando. Alcune delle tecniche classificate in questa tipologia sono il modello di regressione lineare, la regressione logistica, etc.
- *Machine learning*: è un ramo dell'intelligenza artificiale che oltre ad essere utilizzato per sviluppare tecniche per consentire ai computer di apprendere delle conoscenze, viene impiegato in una vasta gamma di settori, come rilevamento di frodi, grazie ad una serie di metodi statistici avanzati per la regressione.

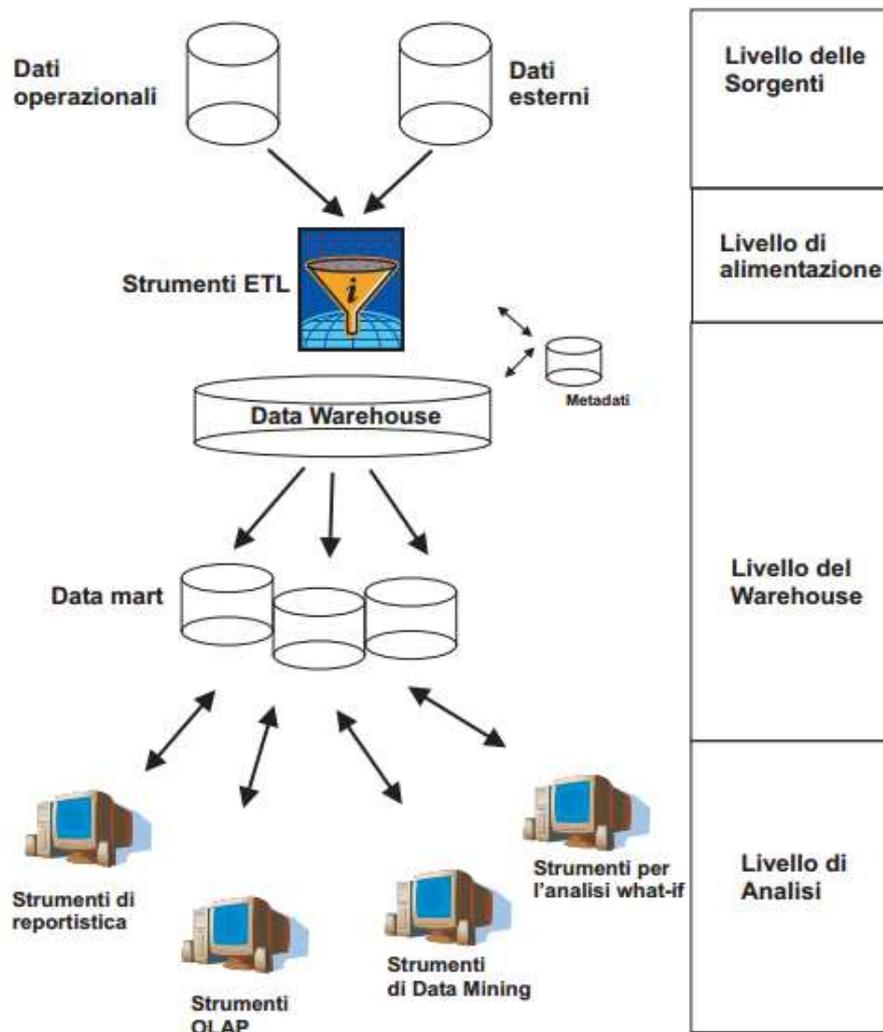


Figura 1.1: Struttura a tre livelli di un DW

## 1.2. I Big Data

Nel 2010, il volume di dati presenti nel cloud era di 500 miliardi di gigabyte (Bollie, 2010). Il Sole 24 Ore nell'articolo di Infodata del 2 Novembre 2015 intitolato 'Big Data, tutti i numeri che contano' prevede che entro il 2020 questo numero sarà destinato ad arrivare a 44 trilioni di gigabyte, cioè ad aumentare di circa 90 milioni di volte. Risulta evidente come la quantità di dati di cui un'azienda può usufruire cresce di giorno in giorno e sarà destinata a raggiungere numeri inimmaginabili. Questo fenomeno, se da un lato è visto come un insieme sconfinato di dati confusi e ridondanti di difficile controllo, dall'altro può essere considerato come una nuova opportunità di generazione di informazioni.

Per capire meglio cosa sono i Big Data è necessario approfondire i concetti sui quali questo fenomeno si basa e dai quali esso ha avuto origine.

- Uno degli aspetti più importanti e più evidenti che caratterizza i Big Data è il *volume* dei dati digitali che è a disposizione delle aziende di tutto il mondo. Ogni giorno, sempre più dati vengono automaticamente generati da persone nell'ambito domestico e aziendale, attraverso smartphone e tablet, da eventi che accadono, dai social network, da nuovi sistemi di acquisizione dati e di salvataggio nel cloud. Di conseguenza, e a differenza del DW, i dati presenti in internet sono altamente eterogenei e inconsistenti. Inoltre non è presente alcuno standard per il caricamento delle informazioni, motivo per cui i Big Data sono caratterizzati da un notevole "rumore".
- Altro aspetto che caratterizza i Big Data è, quindi, la *varietà* dei dati dovuta proprio alla molteplicità delle fonti che possono essere sia interne all'azienda sia esterne.
- Ultima caratteristica è la *velocità* con cui i dati vengono generati e caricati.

Queste tre caratteristiche peculiari dei Big Data (volume, varietà, velocità) rappresentano il Modello delle 3v, creato dal Data Warehouse Institute (TDWI) per cercare di descrivere questo fenomeno. Se si considerano questi tre aspetti assieme si ottiene la quarta "v", quella più importante: il *valore*. Sempre più aziende vedono nell'utilizzo e nell'analisi dei Big Data una modalità per aumentare il numero e il valore delle informazioni ricavabili dai dati. Per le caratteristiche sopracitate, risulta però evidente come questo non sia affatto facile ed economico, in quanto l'enorme quantità di informazioni non può essere contenuta e gestita in un normale RDBMS e portare ad un'analisi rapida ed efficace. Le tecnologie utilizzate, che devono essere altamente scalabili e basate sulla noSQL, risultano di conseguenza molto costose sia nell'installazione che nella gestione. E' da evidenziare, però, che sempre più si stanno diffondendo dei tool (anche open source) che permettono l'analisi dei Big Data attraverso il cloud e che possono essere alla portata anche delle piccole e medie imprese.

A questo punto si può arrivare alla definizione che Pasini e Perego hanno adottato nel rapporto di ricerca intitolato "Big Data: nuove fonti di conoscenza aziendale e nuovi modelli di management", pubblicato su SDA Bocconi nel Dicembre 2012:

*"Big Data è un nuovo concept di conoscenza aziendale degli oggetti e degli eventi di business che fa leva sull'attuale varietà dei dati, in termini di formati (strutturati e non strutturati) e fonti (interne ed esterne), sull'aumentata velocità*

*di generazione, raccolta, aggiornamento ed elaborazione dei dati (in tempo reale, in streaming, dati “tracciati”) e sul crescente volume dei dati stessi, al fine di generare nuove analisi e insights, in ogni settore economico, in imprese di ogni dimensione, precedentemente considerate oltre le capacità tecniche e interpretative disponibili e per scoprire infine un nuovo potenziale valore di business; per ottenere questi risultati le imprese devono gestire gli appropriati fattori abilitanti, di tipo organizzativo, tecnologico e normativo”*

I Big Data, come la BI, hanno l'obiettivo di generare informazioni a partire da un insieme di dati al fine di migliorare il processo di decisione aziendale. Tuttavia, la differenza nelle tipologie e nella varietà dei dati, obbliga una diversa modalità di analisi dei dati. La BI utilizza la statistica descrittiva per analizzare dati ad alta densità di informazione, cioè utilizza dati strutturati e organizzati secondo determinate regole aziendali e di business. Viceversa i Big Data utilizzano la statistica inferenziale, basata su sistemi non lineari che permettono di individuare dipendenze e rapporti e dedurre leggi da grandi insiemi di dati altamente eterogenei e non strutturati.

### 1.2.1. Le PMI e i Big Data

Il fenomeno dei Big Data, già altamente diffuso nel mondo, è in grande aumento anche in Italia, anche se sono ancora presenti delle problematiche soprattutto in termini di governance e competenze. La consapevolezza che la capacità di immagazzinare, conservare e analizzare flussi di dati per ricavarne delle informazioni strutturate permetta di generare un vantaggio in termini di competitività, è sempre più diffusa nelle aziende italiane anche di piccole dimensioni. L'analisi dei Big Data sta diventando uno degli strumenti principali nel management e nell'individuazione delle nuove strategie di business. Non manca, però, anche la consapevolezza che per poter gestire questa tipologia di informazioni è necessaria una figura specializzata in grado di coordinare conoscenze informatiche e statistiche a conoscenze in ambito aziendale e di business. Questa figura, che prende il nome di *data scientist*, non è diffusa nel nostro Paese e per le aziende risulta molto costoso formare il proprio personale. A questo proposito, stanno nascendo delle società che mettono a disposizione delle figure qualificate a prezzi non eccessivi. Inoltre, negli ultimi periodi, sempre più prodotti di BI, inseriscono dei tool nelle loro suite per l'analisi dei Big Data e lo storage dei dati gestiti completamente in cloud. Offrono, cioè, un'*infrastructure as a service* in modo da consentire anche alle piccole e medie imprese di usufruire di questa immensa risorsa senza dover necessariamente possedere un'infrastruttura fisica. I Big Data rappresentano, probabilmente, l'occasione

per le PMI di ridurre il gap presente con le grandi imprese, ma anche di differenziarsi dai diretti concorrenti nel mercato grazie ad una attività di analisi e di previsione, con l'obiettivo di migliorare tutti i processi decisionali e di business. E' presente, tuttavia, ancora un certo scetticismo, dovuto principalmente ad una conoscenza poco approfondita delle potenzialità derivanti dal mondo di Internet e delle occasioni offerte dal mercato dei software BI.



# CAPITOLO 2

## Panoramica del mercato

Per raggiungere i propri obiettivi, un'azienda deve investire su un'infrastruttura solida che le permetta di raccogliere, uniformare, elaborare, analizzare e visualizzare tutti i dati che ha a disposizione. In tutto quest'enorme lavoro, l'information technology manager deve scegliere e gestire i software da utilizzare. Effettuare una scelta corretta degli strumenti da analizzare non è semplice e richiede la conoscenza di una grande quantità di realtà per capire quali prodotti sono in grado di soddisfare i bisogni dell'azienda. Ecco perché, prima di arrivare ad un confronto tra prodotti, è necessario avere una panoramica generale su quello che offre il mercato dei software di business intelligence. Come accennato nell'introduzione, in questa analisi verranno considerati prodotti classificati in tre categorie: i prodotti storici, i prodotti indipendenti e i prodotti open source. Prima di addentrarsi nell'analisi, è utile chiarire un particolare: tra i prodotti storici è stato inserito anche MicroStrategy. In realtà, questo software apparterebbe ai prodotti indipendenti ma, grazie ad una diffusione molto vasta e a dei fatturati in linea con quelli degli altri prodotti storici, può rientrare a pieno diritto tra questi. E' l'esempio di come un prodotto non appartenente alle maggiori società sia riuscito a conquistarsi, grazie a grandi investimenti e a scelte mirate, un posto da leader nel mercato.

### 2.1. Prodotti Storici

La prima tipologia che verrà affrontata è quella dei prodotti storici. Con questo termine si vogliono indicare quei prodotti che sono presenti da molto tempo nel mercato della business intelligence e dell'informatica in generale e ai quali appartengono i fatturati più alti, una maggiore diffusione e un'offerta di prodotti che non si limita ad una soluzione soltanto, ma ad un numero di possibilità che permette di coprire e soddisfare tutte le esigenze di un'azienda. Per questo motivo, più che di prodotti, sarebbe corretto parlare di aziende, in quanto per ognuna di essa verranno analizzati più prodotti diversi. E' necessario sottolineare come spesso queste grandi potenze acquistino prodotti solamente per toglierli dal mercato e creare così una carenza dell'offerta da poter colmare con uno o più dei prodotti sviluppati da queste aziende. Altre volte, invece,

decidono di acquistare un prodotto non tanto per coprire il buco nell'offerta del mercato, quanto per far fronte alle carenze delle loro soluzioni. Questo è il caso, per esempio, di SAP, che ha acquistato Business Objects. Le aziende analizzate in seguito sono quelle che vengono definite le cinque maggiori produttrici di software al mondo: Oracle Corporation, Microsoft Corporation, IBM, SAP e MicroStrategy.

### 2.1.1. Oracle Corporation

Oracle Corporation è una delle più famose società di software nel mondo, fondata nel 1977 con sede in California, nella Silicon Valley, da Larry Ellison, individuato dalla rivista statunitense *Forbes* come il quinto uomo più ricco del mondo nel 2015. Il prodotto di punta è Oracle, uno dei più diffusi database management system (DBMS), la cui prima versione risale al 1977; da allora ha subito una serie di modifiche e aggiornamenti fino ad arrivare alla versione più recente chiamata Oracle Database 12c, che verrà analizzata nel capitolo successivo.

Oltre all'infrastruttura fisica rappresentata dal database, sono necessari dei tool sia per la parte di back end, quindi di caricamento e di integrazione dei dati, sia di front end, cioè di reporting e analisi. Oracle ha investito molto in entrambe le aree, fornendo una quantità importantissima di prodotti e di soluzioni.

Uno dei prodotti principali di Oracle è **Oracle Data Integrator** (ODI), uno dei tool di back end più potenti e completi dal punto di vista delle funzionalità presenti nel mercato. ODI permette di estrarre dati da un numero molto elevato di sorgenti, come database Oracle, Microsoft SQL Server, file CSV, database MySQL, Hyperion, IBM DB2, Microsoft Access, file Excel e, attraverso un'interfaccia abbastanza complessa, ma molto potente, è in grado di elaborare, integrare e certificare i dati da salvare nella struttura di destinazione, che viene chiamata *target*. ODI permette di creare dei modelli grazie ai quali si creano i flussi di lavoro e le procedure per l'integrazione e il caricamento dei dati. La caratteristica che distingue ODI dagli altri strumenti ETL presenti nel mercato è la possibilità, attraverso un meccanismo di contesti, di utilizzare le stesse procedure per caricare i dati su strutture fisiche differenti. E' possibile, cioè, far puntare uno stesso modello logico a più architetture fisiche semplicemente modificando i contesti. Le funzionalità di questo prodotto sono senza dubbio tra le più complete e potenti di tutti i prodotti di ETL ma, di conseguenza, anche la difficoltà di utilizzo e di apprendimento è tra le più elevate.

Grazie ad uno strumento potente come ODI è possibile ottenere una base di dati coerente e consistente sulla quale poter appoggiare i moltissimi prodotti per la business intelligence offerti da Oracle. Tra questi vale la pena soffermarsi su due piattaforme: Oracle Hyperion e Oracle Business Intelligence.

**Oracle Hyperion** è un software front end che offre la possibilità agli utenti, attraverso un'interfaccia intuitiva, di sincronizzare tutte le informazioni aziendali in un ambiente web nel quale sono riproposte le funzioni tipiche di Windows come i menù a barre, drag-and-drop, etc. Offre tutte le funzionalità che ci si aspetta da uno strumenti BI, in particolar modo:

- *Reporting*: grazie ad Oracle Hyperion è possibile creare report sia statici che interattivi su grandi quantità di dati per controllare le performance aziendali. I report possono, di conseguenza, essere pubblicati e distribuiti alle varie aree di interesse come, ad esempio, il controllo di gestione.
- *Dashboard*: Hyperion offre la possibilità di realizzare dei dashboard altamente interattivi grazie ai quali, mediante pochi click, è possibile navigare nei dati per scoprire le cause di determinati eventi. Generalmente, i dashboard vengono utilizzati per monitorare i principali indicatori di performance dell'impresa.
- *Analisi*: oltre alle due funzionalità appena elencate, questo prodotto permette di ricavare informazioni aggiuntive dai dati, analizzandoli ed elaborandoli direttamente in Excel oppure in un ambiente web.

Oracle Hyperion, pur essendo un tool front end, offre la possibilità di estrarre dati da diverse fonti, come database relazionali o multidimensionali o semplici file.

**Oracle Business Intelligence** è un secondo prodotto BI offerto da Oracle Corporation. Anche OBI, come Hyperion, offre una serie di componenti che permettono di soddisfare le esigenze di analisi e di visualizzazione dei dati. I principali tool offerti da questa piattaforma sono:

- *Oracle Business Intelligence Answers*: offre capacità di interrogazioni e analisi ad-hoc. Consiste in un ambiente puramente web progettato sopra ad uno strato logico in cui la complessità delle informazioni viene nascosta e rielaborata, in modo da permettere all'utente di interfacciarsi con una struttura semplificata dei dati.
- *Oracle Business Intelligence Interactive Dashboard*: grazie a questo componente è possibile realizzare in modo grafico dei dashboard altamente interattivi, sempre

attraverso un ambiente web. Permette di inserire all'interno di questi elaborati tabelle, grafici, report e molto altro.

- *Oracle Business Intelligence Publisher*: motore di reporting altamente scalabile che permette di realizzare report con dati provenienti da diverse sorgenti e pubblicarli in vari formati a seconda delle esigenze.

A questo punto risulta utile analizzare quelle che sono le differenze tra le due alternative front end offerte da Oracle Corporation. Innanzitutto Hyperion si basa esclusivamente su database relazionale, in quanto comprende un motore ROLAP, mentre Oracle BI supporta analisi di tipo multidimensionali e, di conseguenza, tempi di risposta sostanzialmente immediati. Altra caratteristica che differenzia Hyperion è la possibilità di eseguire l'analisi dei dati direttamente su Microsoft Excel. Questo aspetto è fondamentale per molte aziende perché consente di utilizzare funzionalità avanzate attraverso un software semplice e di utilizzo comune. D'altra parte però, Oracle BI mette a disposizione degli utenti una serie di strutture predefinite sia per i report che per i dashboard, semplificandone notevolmente la realizzazione e l'implementazione. Di conseguenza, ne risulta che Oracle BI è più semplice da utilizzare rispetto ad Hyperion, che viene considerato più tecnico, ma con delle funzionalità, soprattutto in ambito di analisi predittiva, maggiori e più efficienti. La scelta fra le due piattaforme dipende quasi esclusivamente dalle esigenze delle singole aziende.

Oracle è uno dei prodotti più diffusi e più importanti sia nel settore delle infrastrutture database, sia in quello dei software di BI. Uno degli aspetti più importanti di Oracle, e il principale motivo secondo Gartner per cui le aziende lo scelgono, è l'elevato grado di integrazione dei prodotti con le varie applicazioni aziendali, assieme alla copertura pressoché totale di tutti i settori di un'impresa e, di conseguenza, delle relative esigenze. Prodotto che sicuramente spicca su tutti, al di là del database Oracle, è Oracle Data Integrator, uno dei più potenti e completi strumenti ETL offerti dal mercato. Questa potenza, però, comporta un'elevata complessità del prodotto che rende difficile l'apprendimento e, di conseguenza, l'autonomia dell'utente. Si tratta di una complessità che caratterizza più o meno tutti i prodotti, assieme ad un costo che non sempre le medie imprese possono permettersi. Infine, per quanto riguarda l'architettura cloud e l'analisi, Oracle sta investendo molto, anche se è ancora in leggero ritardo rispetto agli altri prodotti storici analizzati. Sicuramente la nuova tecnologia database permetterà ad Oracle di recuperare il terreno perso in questo ambito e tornerà ai massimi livelli anche sotto questo punto di vista.

### 2.1.2. Microsoft Corporation

Microsoft Corporation è la più importante produttrice di software al mondo per quanto riguarda il fatturato ed è una delle più importanti aziende di informatica presenti nel mercato internazionale. E' stata fondata nel 1975 da Paul Allen e Bill Gates con sede a Redmond, nello stato di Washington, con il nome di "Micro-Soft Company". Microsoft è azienda leader nel mercato grazie ad alcuni suoi prodotti come il sistema operativo Microsoft Windows, la suite Microsoft Office, dispositivi hardware come console di gioco Xbox, sistema operativo mobile Windows Phone e molto altro. MS è sicuramente una delle dirette concorrenti di Oracle grazie a Microsoft SQL Server, una tecnologia database attorno alla quale ha costruito e sviluppato tutte le soluzioni di DW e BI. Tutti i tool per la BI sono delle funzionalità che si eseguono attraverso uno strumento chiamato Visual Studio, un ambiente di sviluppo creato da Microsoft in grado di supportare diversi linguaggi di programmazione, tra cui C, C++, C#, Visual Basic, etc. In particolar modo, le principali funzionalità supportate dalla versione Business Intelligence di Microsoft SQL Server 2014 sono:

- **SQL Server Integration Services:** è una piattaforma che permette di compilare soluzioni per l'estrazione, l'integrazione e il caricamento dei dati a livello aziendale. Tutti i flussi di lavoro sono suddivisi in pacchetti, creati a partire da una serie di step già esistenti in forma grafica, in modo da semplificare il lavoro dell'utente. Inoltre, Integration Services offre la possibilità di creare delle operazioni personalizzate attraverso la scrittura di codice Java o SQL. Grazie a questo strumento si possono estrarre file da una vasta gamma di sorgenti diverse, come file di dati XML, database relazionali e file flat, per poi caricarli in una o più destinazioni dei dati. Integration Services è uno strumento molto potente, nonostante sia di facile utilizzo e apprendimento, caratteristica non comune agli strumenti dei principali prodotti presenti nel mercato.
- **Microsoft BI Semantic Model:** è un nuovo concetto introdotto a partire da Microsoft SQL Server 2012 e che, nella versione 2014, si è affermato ulteriormente. Si riferisce al modo in cui Analysis Services e PowerPivot possono essere utilizzati come strato semantico posto sopra ai dati ai fini di reporting. BI Semantic Model offre la possibilità di rappresentare, e di conseguenza di analizzare, i dati aziendali in base ai due principali modelli: relazionale (tabellare) e multidimensionale. Il *Tabular model* si basa su dati strutturati appunto in tabelle, salvate in un database in memory chiamato VertiPaq ed è stato introdotto nella versione di SQL Server del 2012. Esso consente di accedere ai dati in maniera

rapida per renderli disponibili a Excel, lo strumento Microsoft utilizzato per il reporting e l'analisi. Il *Multidimensional model*, invece, si basa sui cubi di dati tipici del modello multidimensionale che è presente all'interno dell'offerta di Microsoft da molto più tempo rispetto al Tabular model. Partendo poi da questo strato semantico, l'utente può eseguire analisi OLAP sia multidimensionali che relazionali. Essendo uno strumento presente da più tempo, il Multidimensional model risulta essere più completo e consolidato, anche se il Tabular model è uno strumento più flessibile grazie all'utilizzo di Excel. **Microsoft Excel** è uno strumento sviluppato da Microsoft e posto all'interno della suite di Office. Esso è uno dei prodotti di front end più diffusi e anche più flessibili tra quelli presenti nel mercato, permettendo inoltre di navigare attraverso i dati.

- **SQL Server Reporting Services:** è uno strumento completamente integrato con i componenti di SQL Server e funziona all'interno di Visual Studio; fornisce una serie di funzionalità per la creazione e la distribuzione di report aziendali. Offre una serie di oggetti predefiniti, come grafici e tabelle, oltre alla possibilità di crearne altri attraverso la programmazione. Grazie a questo prodotto, è possibile integrare all'interno di uno stesso elaborato dati provenienti da diverse fonti e visualizzarli a vari livelli di dettaglio. Reporting Services si basa sullo strato semantico creato attraverso il Multidimensional model e consente agli utenti di schedulare internamente i report creati, di esportarli in vari formati (anche in Excel) e inviarli via email.

Storicamente, Microsoft ha avuto come punto debole la parte di front end e, in particolar modo, l'analisi dei dati. Il prodotto utilizzato generalmente per le soluzioni Microsoft è Excel che, nonostante una grande flessibilità, presenta notevoli limiti nella navigazione in profondità dei dati e anche nella loro visualizzazione. Per far fronte a questa carenza, Microsoft ha sviluppato un concetto che negli ultimi anni è stato il centro di tutti gli investimenti, sia in termini monetari che di sviluppo. Questo concetto è ciò che viene definito "*personal BI*" e che indica la possibilità da parte degli utenti di gestire completamente in autonomia tutta la parte relativa alla business intelligence. Grazie a questa mentalità, si è arrivati allo sviluppo di una nuova piattaforma: **Microsoft Power BI**. Per la versione del 2014 di SQL Server, questa piattaforma è stata sviluppata come un ambiente web che presumibilmente verrà integrato all'interno di SQL Server Reporting Services con Microsoft SQL Server 2016. La piattaforma offre funzionalità sia di reporting che di dashboarding, grazie alle quali è possibile creare degli elaborati molto più avanzati e meglio definiti dal punto di vista grafico. Altro aspetto molto importante è la velocità con cui è possibile creare report e dashboard complessi in maniera molto

semplice e intuitiva. Questa velocità è dovuta al fatto che Power BI interroga praticamente in real time il database, aggiornando molto frequentemente i dati e le informazioni visualizzati. In questo modo gli utenti sono in grado di tenere sotto controllo tutti gli andamenti dell'azienda e i suoi parametri. Questa suite supporta anche la parte mobile, che è ben sviluppata e integrata con il resto delle funzionalità, permettendo di visualizzare tutto quello creato via pc direttamente nello smartphone, sia esso Apple, Android o Windows Phone.

Secondo l'analisi e le indagini effettuate da Gartner, Microsoft tende ad essere ancora sottovalutata dalle imprese per quanto riguarda le soluzioni enterprise offerte dall'azienda. D'altro canto, è una delle soluzioni più adottate e apprezzate per le piccole e medie imprese. In realtà, Microsoft offre una serie di prodotti che hanno nella loro integrazione con SQL Server e tra di essi il loro punto di forza. Inoltre, con Power BI Microsoft sta cercando di coprire quelle aree della business intelligence in cui è stata storicamente più debole. Alcune sue funzionalità non sono ancora all'altezza di quelle offerte dai maggiori esponenti del mercato della BI, ma è anche vero che si posiziona in una fascia di prezzo leggermente inferiore ad essi e permette agli utenti di essere autonomi grazie alle interfacce graficamente semplificate.

### 2.1.3. IBM

IBM (*International Business Machine Corporation*) è una delle principali aziende del settore informatico nel mondo. E' stata fondata nel 1911 ad Armonk, New York, anche se era già operativa da alcuni anni. IBM produce e commercializza tecnologie hardware e software, offre infrastrutture, servizi e consulenze in tutto il mondo. Offre soluzioni che vanno a soddisfare praticamente ogni esigenza delle imprese, come l'integrazione dei dati, business management, data warehousing, cloud computing, sicurezza e molto altro. Tra tutti i prodotti offerti da IBM, l'analisi di questo lavoro si soffermerà su tre prodotti: uno di back end per l'integrazione dei dati e due di front end: uno per l'analisi e il reporting e l'altro per l'analisi predittiva. Queste tre soluzioni sono rispettivamente *IBM InfoSphere DataStage*, *IBM Cognos* e *IBM SPSS*.

Nel sito ufficiale di IBM (<http://www.ibm.com/>) **IBM InfoSphere DataStage** viene presentato come un prodotto in grado di integrare i dati di più sistemi offrendo anche la versione enterprise e la gestione dei metadati. La piattaforma ha la caratteristica di essere scalabile, fornendo l'integrazione di tutti i tipi di dati, compresi i big data. In particolar modo InfoSphere DataStage offre le seguenti funzionalità:

- *Piattaforma ETL scalabile*: supporta la raccolta, l'integrazione e la trasformazione di volumi elevati di dati, aventi strutture anche complesse. Altra funzionalità molto importante all'interno di uno strumento di back end è la possibilità di integrare i metadati al fine di mantenere una coerenza analitica degli stessi.
- *Supporto per i big data*: consente l'accesso diretto ai big data.
- *Utilizzo semplificato*: consente di aumentare la velocità e l'efficienza di sviluppo e implementazioni dei processi di integrazione attraverso un'interfaccia user-friendly che permette, oltre all'integrazione dei dati, anche la loro validazione.
- *Connettività*: DataStage supporta un numero molto elevato di fonti e di destinazione dati come file di testo, strutture dati XLM, sistemi gestionali (SAP, Oracle...), quasi tutti i database (Oracle, IBM DB2 Universal Database, Microsoft SQL Server...), Web Service, etc.

Il punto di forza di IBM è sicuramente Cognos 10, una piattaforma che offre funzionalità di BI, analisi, gestione delle strategie e delle prestazioni aziendali e che permette alle medie e grandi aziende di migliorare il processo decisionale e ottimizzare i risultati presenti e futuri. Questa piattaforma offre una serie di tool in grado di coprire tutte le funzionalità front end di cui necessita un'azienda. I prodotti disponibili sono integrati all'interno della piattaforma, il che significa che un'azienda può decidere di utilizzare solo alcune delle funzionalità offerte, avendo comunque la consapevolezza di poter ampliare la piattaforma in base alle esigenze che man mano si presentano. La famiglia Cognos comprende tre soluzioni in base alle dimensioni, e quindi alle esigenze, dell'impresa: IBM Cognos Insight, IBM Cognos Express e IBM Cognos Enterprise, che comprende IBM Cognos TM1 e IBM Cognos Business Intelligence.

- **IBM Cognos Insight**: soluzione di analisi personale in grado di creare dashboard e applicazioni personalizzate sul desktop. Grazie a questo programma, gli utenti possono importare i dati di vario tipo (file personali, database...), integrarli e analizzarli per poi sviluppare dashboard per la visualizzazione grafica dei risultati ottenuti. Cognos Insight comprende, inoltre, un insieme di regole avanzate e la possibilità di crearne di personalizzate, che possono essere utilizzate per la modellazione di scenari "What-if" in grado di eseguire delle simulazioni e quindi delle previsioni su più ambiti aziendali.
- **IBM Cognos Express**: fornisce tutte le caratteristiche elencate anche in precedenza (analisi, dashboarding, pianificazione, previsione...) con l'aggiunta di altre funzionalità, come la possibilità di presentare i dati attraverso dei report,

l'integrazione con Microsoft Office e il supporto mobile per iPad. Questo prodotto rappresenta una soluzione scalabile utilizzata prevalentemente dalle imprese di medie dimensioni, in quanto permette di ridurre leggermente i costi.

- **IBM Cognos TM1:** software di pianificazione enterprise che permette di trasformare l'intero ciclo di pianificazione e analisi. E' disponibile sia in versione client che in cloud, offrendo anche delle soluzioni mobile. Grazie alle procedure guidate, Cognos TM1 permette di creare dei modelli di analisi e di pianificazione in modo semplificato e offre la possibilità di monitorare le metriche per misurare i progressi degli obiettivi. Da queste misurazioni è possibile allineare quelle che sono le decisioni, e poi le azioni in ambito di business. TM1 permette, inoltre, di analizzare i grandi dataset che vengono modificati frequentemente e di organizzarli in modelli complessi utilizzando un motore OLAP multidimensionale in memory.
- **IBM Cognos Business Intelligence:** offre una soluzione completa che ha l'obiettivo di coprire tutti gli ambiti in cui è possibile eseguire un'analisi per migliorare il business dell'impresa. Offre funzionalità mobile non solo per Apple, ma anche per Android e altri dispositivi, attraverso i quali è possibile visualizzare e modificare dashboard e report, progettati in modo tale da essere integrati tra di loro. Questa soluzione permette anche l'analisi dei big data e dei dati sulla gestione delle prestazioni in un formato di foglio elettronico. Cognos Business Intelligence cerca di integrare tutte le funzionalità di BI in un'unica soluzione, offrendo inoltre un kit per lo sviluppo software.

Key capabilities	Cognos Insight	Cognos Express	Cognos Enterprise
Dashboards	X	X	X
Analysis	X	X	X
What-if scenario modeling	X	X	X
Planning and budgeting	X	X	X
Production reports		X	X
Microsoft Office integration		X	X
Support for Apple iPad		X	X
Support for other mobile devices			X
Enterprise-wide collaboration			X
Software Development Kit			X

Tabella 2.1. Comparazione delle principali funzionalità di Cognos Insight, Cognos Express e Cognos Enterprise (Cognos TM1 assieme a Cognos Business Intelligence). Fonte: IBM Knowledge Center.

Un ruolo molto importante all'interno della grande famiglia di prodotti forniti da IBM ce l'ha sicuramente IBM SPSS, un prodotto che va a coprire tutte le esigenze di analisi predittiva, scoperta di nuove informazioni e previsioni future. Comprende principalmente due strumenti:

- **IBM SPSS Statistics Base:** è un software di analisi statistica che fornisce le funzionalità fondamentali necessarie per gestire l'intero processo analitico, al fine di scoprire e ricavare nuove informazioni da quelle che si hanno a disposizione. Comprende una serie completa di procedure statistiche (tabulazioni incrociate, frequenze, medie, regressione lineare, tecniche di simulazione Monte Carlo...) e di tecniche integrate che hanno l'obiettivo di ristrutturare i file prima dell'analisi. SPSS Statistics Base supporta anche la visualizzazione avanzata dei risultati delle analisi attraverso report OLAP, grafici e diagrammi, anche su dispositivi mobile.
- **IBM SPSS Modeler:** è una piattaforma di analisi predittiva progettata per introdurre l'intelligence predittiva nel processo decisionale gestito da persone. Attraverso lo sviluppo di modelli predittivi con una vasta gamma di algoritmi avanzati, integrati con regole di business e tecniche di ottimizzazione, ha lo scopo di offrire dei suggerimenti a persone e sistemi responsabili dei processi decisionali al fine di semplificarne tutte le operazioni. Sono disponibili diverse versioni di cui una anche in cloud.

Uno degli aspetti positivi di IBM è sicuramente Cognos, prodotto molto diffuso e completo. Sono da evidenziare soprattutto le funzionalità di reporting, dashboarding, gestione dei metadati e tutta la parte di analisi (What-if e OLAP). Offre una serie molto vasta di funzioni e tool per la gestione di tutta l'impresa; infatti è principalmente utilizzata dalle grandi aziende per installazioni di dimensioni importanti.

Aspetti che non convincono del tutto sono la complessità dei programmi, nonostante l'utilizzo di interfacce grafiche, e la gestione delle licenze, che rende difficile la stima di un eventuale investimento da fare da parte di un'azienda per implementare una soluzione basata su prodotti IBM. Infine, sempre secondo Gartner, i clienti tendono a rimanere insoddisfatti dall'esperienza personale effettuata con questo prodotto. Con 'esperienza personale' vengono intesi una serie di parametri come l'assistenza, aggiornamenti e migrazioni, facilità di utilizzo e, di conseguenza, disponibilità di conoscenze.

Nonostante queste note non propriamente positive, rimane uno dei prodotti migliori e più consolidati nel mercato grazie anche alla nuova versione di Cognos (*IBM Cognos 10*).

#### 2.1.4. SAP

SAP SE (*Systeme, Anwendungen, Producte in der Datenverarbeitung*, cioè *Sistemi, Applicazioni e Prodotti nell'elaborazione dati*) è una multinazionale europea per la produzione di software, attiva nel mercato dal 1972 quando fu fondata a Walldorf, in Germania. E' una delle principali aziende del mondo nel settore degli ERP, cioè dei software gestionali; i prodotti offerti in questo settore sono SAP R/3 e SAP Business One. SAP non si è concentrata esclusivamente su software gestionali, ma ha prestato molta attenzione a temi come data warehouse e business intelligence.

Un prodotto fondamentale in questo ambito è sicuramente **SAP Business Warehouse** (BW) che rappresenta la più importante soluzione enterprise di DW offerta da SAP. BW è una combinazione di DB e strumenti di gestione di database che vengono utilizzati per supportare la gestione dei processi decisionali. E' un prodotto che è supportato sia da RDBMS standard che da SAP HANA, cioè un RDBMS (*relational database management system*) in-memory e column-oriented sviluppato da SAP. Oltre al database, BW offre anche degli strumenti per l'estrazione dei dati, l'analisi e il reporting. Permette di estrarre ed integrare dati provenienti sia dalle applicazioni SAP che da altre fonti, anche esterne all'azienda, al fine di trasformare e consolidare le informazioni aziendali. Questo programma fornisce un'infrastruttura ad alte prestazioni che permette di analizzare e gestire le informazioni in modo rapido ed efficace. Risulta evidente come sia necessario avere una struttura fisica di back end, cioè una base di dati coerente e ben strutturata, affinché sia possibile usufruire di tutti gli strumenti di front end offerti da SAP BW. Altre caratteristiche di questo prodotto sono: la presenza di un motore OLAP integrato per l'analisi multidimensionale, il supporto di più lingue, un repository per i metadati e la presenza di Business Explorer, un'interfaccia utente basata sul web.

Recentemente SAP ha investito molto anche nella parte prettamente di front end, acquistando nel 2007 Business Objects, una società specializzata nella realizzazione di software per l'analisi e il reporting, dalla quale ha rilevato l'omonimo software. SAP ha inserito Business Objects all'interno di una suite, chiamata **SAP Business Objects Business intelligence** in cui sono stati integrati altri prodotti, tra cui anche uno per la parte di estrazione e integrazione dei dati. Tra tutti i tool presenti all'interno di SAP BO BI risulta importante, ai fini di questo lavoro, soffermarsi su uno strumento di front end (SAP Business Objects Web Intelligence) e uno di back end (SAP Data Services).

- **SAP Business Objects Web Intelligence:** è un tool fortemente orientato al reporting che permette, infatti, di realizzare report ad-hoc per soddisfare tutte le esigenze di visualizzazione e di stampa dei dati da parte di tipologie di utenti

differenti, che vanno dalla direzione della produzione al controllo di gestione, fino ad arrivare alla dirigenza. Esistono due versioni di Business Objects Web Intelligence: una consiste in un'applicazione desktop, mentre l'altra in una soluzione via browser. La caratteristica principale di questo prodotto è la presenza di un tool per la creazione e la gestione degli universi che consistono in una rielaborazione e una rappresentazione logica dei dati. Questo strato logico su cui si basa Business Objects Web Intelligence è fondamentale in quanto permette agli utenti che devono realizzare dei report di non interfacciarsi direttamente con la struttura fisica del database e quindi delle tabelle. Grazie agli universi è inoltre possibile stabilire determinate regole e variabili, creare tabelle derivate e suddividere i dati in base al contesto aziendale. L'analisi è probabilmente l'aspetto più debole di questo tool, ma anche, in generale, di tutta la piattaforma SAP BO BI. Business Objects è un buon prodotto che permette di realizzare report, contenenti oggetti come tabelle, testi e grafici, attraverso un'interfaccia basata sulla tecnologia *drag-and-drop* e un layout orientato fortemente alla stampa del report. Inoltre, tutti gli elaborati possono essere schedulati internamente a BO ed esportati sia come report che in formato pdf e xls.

- **SAP Data Services:** è un software ETL per l'integrazione e la trasformazione dei dati. Permette di estrarre i dati da sorgenti differenti (applicazioni, servizi web, database, ecc), quindi di elaborarli trasformandoli e combinandoli, per poi caricarli nelle stesse sorgenti di partenza o in un target comune come per esempio un database. SAP Data Services offre tutte le funzionalità standard di trasformazione dei dati di un tool di back end: rappresentazione delle relazioni tra le tabelle attraverso le operazioni di join, ricerche, filtri, aggregazioni e calcoli. Presenta inoltre una serie di strumenti per l'analisi del testo, controllo dei dati e supporto delle operazioni di *data quality*. Questo prodotto si integra molto bene sia con Business Objects Web Intelligence che con SAP BW.

SAP offre due soluzioni pensate principalmente per le piccole imprese: SAP Crystal Report e SAP Crystal Dashboard. Specialmente Crystal Dashboard va a coprire la parte di BI che consiste nella visualizzazione ed esplorazione dei dati, anche se le potenzialità non sono propriamente all'altezza di altre soluzioni offerte dal mercato. Nonostante ciò, offre comunque delle funzionalità interessanti, come la possibilità di esportare gli elaborati per esempio in pdf.

Tra gli aspetti più positivi di SAP spicca sicuramente il grande impegno e l'elevata quantità di investimenti al fine di aumentare le potenzialità dei prodotti in modo da soddisfare tutte le esigenze aziendali. SAP infatti offre soluzioni enterprise che vanno dai software gestionali fino a prodotti front end per la BI. Un ruolo molto importante ce l'ha la piattaforma SAP Business Objects Business Intelligence che copre sostanzialmente tutti gli ambiti BI. L'elemento più debole, d'altro canto, è senza dubbio il dashboarding; se il reporting è molto sviluppato e consolidato grazie a Business Objects Web Intelligence, la parte di creazione dei cruscotti interattivi è carente e non all'altezza dei prodotti appartenenti alla fascia di mercato più alta. Altro aspetto negativo è quello che Gartner, nella sua analisi, definisce "esperienza dell'utente". Nonostante il costante impegno di SAP nella semplificazione delle soluzioni e delle licenze, la qualità dei prodotti è ritenuta dagli utenti non all'altezza di altri concorrenti nel mercato dei prodotti di BI. A prova di ciò vanno considerate, per esempio, la migrazione dei report dalla versione 3 di BO alla versione 4 che sostanzialmente è impossibile o la difficoltà da parte degli utenti ad utilizzare questi prodotti.

SAP è leader indiscusso nel campo dei software gestionali e sta investendo molto per affermarsi a pieno merito anche nel campo dei prodotti BI. E' già in grado di fornire agli utenti una serie di prodotti per ogni tipo di esigenza, inoltre, ha cominciato ad offrire soluzioni molto interessanti in cloud e mobile. E' un prodotto adottato generalmente da grandi imprese che sfruttano l'integrazione dei prodotti in tutte le aree dell'impresa.

#### 2.1.5. MicroStrategy

MicroStrategy Inc. è un'azienda fornitrice di software di business intelligence, fondata nel 1989 da Michael J. Saylor e Sanju Bansal. Ha sede nell'area metropolitana di Washington DC e possiede filiali in 26 paesi del mondo, tra cui anche l'Italia, a Milano. Grazie a tutti questi anni di esperienza, MicroStrategy è arrivata a sviluppare una delle piattaforme per la BI più complete per quanto riguarda le funzionalità, che riescono a coprire sostanzialmente ogni richiesta e necessità delle aziende. E' una soluzione utilizzata soprattutto dalle grandi imprese per implementare sistemi IT molto complessi assieme a soluzioni di data warehousing.

**MicroStrategy Intelligence Server** è un server BI che permette alle aziende di integrare tutte le applicazioni front end in un'unica soluzione back end. Grazie a MicroStrategy Intelligence Server le imprese possono standardizzare in un'unica tecnologia tutti i dati e di conseguenza tutte le funzionalità (reporting, analisi, dashboarding, OLAP visualization e data discovery), permettendo agli utenti di accedere a tutte le applicazioni

di BI. Le caratteristiche principali di questa soluzione sono l'unicità della base dati grazie alla quale è possibile avere un front end certificato, maggior facilità di gestione di tutta l'infrastruttura e la possibilità da parte degli utenti di accedere a tutti i dati a disposizione dell'azienda. Questi sono tutti i vantaggi ottenuti grazie all'adozione da parte di MicroStrategy di un'architettura BI di tipo enterprise e non dipartimentale, che permette di ottenere prestazioni migliori a costi più contenuti. Risulta quindi evidente come MicroStrategy abbia fatto propria l'architettura a tre livelli e sia riuscita a ricavarne tutti i vantaggi che ne derivano. In questo tipo di architetture è possibile fornire agli utenti una base dati che sia il risultato di una serie di dati integrati a partire da sorgenti e tecnologie diverse. In particolar modo MicroStrategy Intelligence Server permette di estrarre dati da tutti i principali database, ad esempio IBM DB2, Microsoft SQL Server, MySQL, Teradata, Oracle, etc. Tuttavia MicroStrategy non offre un tool di ETL per determinare il flusso di estrazione e caricamento dei dati. Questa mancanza, però, non comporta un problema così grave in quanto, avendo un'infrastruttura ben organizzata, è possibile associarci uno strumento ETL di altri vendor, ottenendo così una soluzione completa.

Un prodotto molto importante all'interno della suite è **MicroStrategy Architect**, uno strumento di sviluppo rapido che mappa la struttura fisica di uno o più sorgenti di dati in un modello logico object-oriented del business. Fornisce un'interfaccia grafica e delle procedure guidate per definire le regole di business, per la creazione facilitata e il mantenimento delle applicazioni BI sviluppate all'interno di un ambiente uniforme. Grazie a queste interfacce è possibile creare un'astrazione logica della struttura fisica dei database di origine, inserendo e togliendo le tabelle in modo grafico e definendo le associazioni fra le tabelle in modo da riprodurre il più fedelmente possibile le relazioni presenti tra le varie entità (uno-a-uno, uno-a-molti e molti-a-molti). Tutte queste informazioni sono poi memorizzate all'interno del repository centralizzato dei metadati.

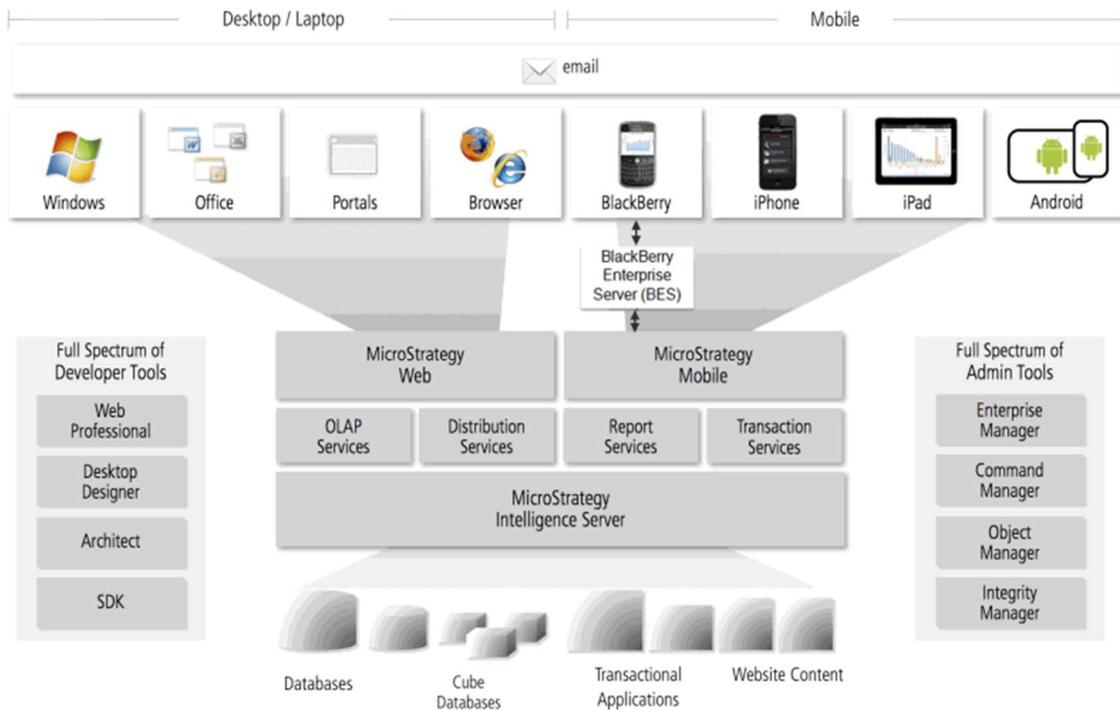


Figura 2.1 Architettura della piattaforma MicroStrategy. (Fonte: sito ufficiale MicroStrategy ([www2.microstrategy.com](http://www2.microstrategy.com)))

Anche per quanto riguarda il front end, MicroStrategy offre una serie di prodotti molto validi che vanno a ricoprire quasi tutte le funzionalità che ci si aspetta a livello di BI. Questi prodotti sono:

- **MicroStrategy Desktop:** è una potente interfaccia di sviluppo che comprende monitoraggio integrato, reporting e analisi via desktop, supportato sia da Windows che da iOS. Grazie a questo prodotto gli utenti possono estrarre dati da sorgenti differenti e salvarli in locale; in questo caso si parla di architettura a due livelli secondo cui le applicazioni accedono direttamente dalle sorgenti. Questo software permette di creare via desktop report aziendali e dashboard all'interno dei quali, grazie alla possibilità di creare filtri, somme, data mining, prompt e molto altro ancora, vengono visualizzati i dati in modo da semplificarne la lettura da parte degli utenti. MicroStrategy Desktop supporta inoltre indagini interattive e analisi OLAP, quindi multidimensionali.
- **MicroStrategy Web:** tool front end che offre sostanzialmente le stesse funzionalità di MicroStrategy Desktop. MicroStrategy ha sempre investito molto nell'architettura web. Infatti, grazie a molti anni di esperienza, è arrivata ad offrire un prodotto scalabile molto completo che presenta molte funzionalità come

reporting, analisi, dashboarding. Inoltre, permette di condividere in tempi brevi tutti i file con tutti gli utenti appartenenti all'impresa. Tra le funzionalità principali ci sono le analisi multidimensionali, il data mining e l'analisi predittiva. MicroStrategy sta sempre più spingendo sul web e sul cloud per offrire soluzioni sempre più scalabili e su misura per ogni tipo di azienda ed esigenza.

- **MicroStrategy Mobile:** MicroStrategy è stata una delle prime aziende a credere e, di conseguenza, a investire molto sul mobile. Ad oggi offre dei supporti per dispositivi diversi come Android, iPad, iPhone e BlackBerry. Grazie a questo prodotto è possibile condividere istantaneamente report e dashboard tra tutti i supporti touch dell'azienda e quindi tener informati tutti gli utenti dando loro anche la possibilità di eseguire delle analisi direttamente dal loro smartphone.

Come già evidenziato precedentemente, i punti di forza di MicroStrategy sono la presenza di una piattaforma integrata e centralizzata alla quale sono collegate tutte le applicazioni per la BI. Inoltre, secondo l'analisi effettuata da Gartner, la maggior parte degli utenti che utilizzano questa soluzione rimane soddisfatta per quanto riguarda la completezza delle funzionalità, la presenza di un repository dei metadati e del supporto sia mobile che web. Certamente non mancano alcuni aspetti negativi. In particolar modo le due problematiche maggiori sono la difficoltà di utilizzo e, di conseguenza, di apprendimento per l'utente, comune a quasi tutti i più importanti prodotti, e la mancanza di uno strumento ETL. Inoltre la migrazione da una versione a quella successiva si è sempre dimostrata complessa e non sempre funzionale.

MicroStrategy si posiziona con pieno diritto tra i prodotti storici sia per quanto riguarda le funzionalità sia, purtroppo, per quanto riguarda i costi. Infatti è una soluzione adottata principalmente da grandi imprese che hanno scelto di realizzare un'infrastruttura che, basata sulla centralizzazione del server e su un sistema di data warehousing, permette di costruire poi un front end molto efficiente.

## 2.2. Prodotti Indipendenti

I prodotti definiti "*Independent*" sono quei prodotti che, nel mercato dei tool di BI, si trovano esattamente in mezzo tra i prodotti storici e quelli open source. Questi tool vengono definiti "indipendenti" in quanto non sono una delle numerose soluzioni offerte dalle maggiori società, ma sono frutto di investimenti da parte di società minori che cercano di farsi spazio all'interno dell'ampia offerta di prodotti grazie alla specializzazione. Si trovano a dover sopravvivere all'interno di una realtà altamente

complessa e piena di concorrenti. Da un lato si trovano a dover competere con le prestazioni e l'integrazione delle funzionalità dei principali prodotti, dall'altro a dover offrire delle soluzioni ad un prezzo competitivo affinché le aziende non preferiscano delle soluzioni open source. Per riuscire ad avere degli sbocchi in questa "giungla", i prodotti indipendenti sono costretti a concentrare sforzi e investimenti in un particolare ambito della BI, che nella maggior parte dei casi si tratta della parte di front end e, in particolar modo, della data visualization. Questo comporta un'integrazione e una copertura delle funzionalità di BI inferiore a quella offerta, per esempio, da Oracle o da SAP ma, d'altro canto, permette loro di offrire una soluzione in parte migliore di quella offerta da tutti gli altri competitors. I prodotti che verranno analizzati in questo paragrafo sono: Qlik, Spotfire e Tableau.

### 2.2.1. Qlik

Qlik è un'impresa di software statunitense con sede a Radnor, in Pennsylvania, anche se originariamente fu fondata nel 1993 in Svezia. Lo sviluppo e la diffusione di questa azienda e dei suoi prodotti sono stati tra i più rapidi nell'ambito della BI. I sistemi tradizionali di BI tendono a ridurre l'agilità dei propri prodotti, e di conseguenza dell'analisi dei dati, per soddisfare le esigenze sempre più complesse delle aziende. A differenza di tutti gli altri prodotti di BI, Qlik ha sviluppato e successivamente brevettato un sistema di analisi in-memory basato sulla teoria associativa. Ciò significa che Qlik permette di caricare i dati in memoria RAM in un database associativo che consente di evidenziare immediatamente le relazioni tra i dati per aiutare ulteriormente il processo decisionale. Questo sistema comporta un accesso molto più rapido ai dati e analisi completamente diverse da quelle basate su database relazionali, in quanto presenta una struttura molto più elastica, basata non tanto sui record quanto più sulle associazioni, e quindi sulle relazioni esistenti tra le varie entità.

Il prodotto principale di Qlik, grazie al quale l'azienda si è diffusa ed è arrivata al successo attuale, è **QlikView**. QlikView è leader mondiale nelle tecniche di analisi in-memory ed è riconosciuta come una delle piattaforme con il più rapido tasso di diffusione al mondo. Questo prodotto ha come indiscusso punto di forza il dashboarding, caratterizzato da un'elevata interattività e una velocità molto significativa. D'altro canto, però, è un prodotto che non supporta, o supporta in maniera limitata, le altre funzionalità che ci si aspetta da uno strumento di BI completo, come il reporting e tutta la fase di back end. QlikView permette di creare dashboard molto complessi e graficamente completi e ricchi di funzionalità, che offrono all'utente finale la possibilità di eseguire delle analisi in maniera

autonoma. L'utente può gestire l'intera analisi e navigare nei dati impostando con semplici passaggi i filtri e selezionando solamente le informazioni di interesse. Questo prodotto offre sostanzialmente un modo estremamente rapido e flessibile per monitorare le prestazioni e i parametri aziendali e tenere sempre controllato l'andamento del mercato e delle vendite attraverso la visualizzazione grafica di tutte le informazioni. Fra gli oggetti grafici supportati da QlikView ci sono tabelle, elenchi a tendina, grafici anche molto complessi e dinamici, pulsanti e indicatori. Una caratteristica di questo prodotto, come già accennato, è la facilità con cui un utente può navigare nei dati al fine di ricavarne le informazioni di interesse; d'altra parte, però, la creazione del dashboard, seppur si possa effettuare per via grafica, richiede un livello di preparazione da parte dello sviluppatore superiore ad altri prodotti front end presenti nel mercato.

Per far fronte all'esigenza da parte delle aziende di coprire anche le altre aree della BI, Qlik ha lanciato nel mercato, nel settembre del 2014, la piattaforma **Qlik Sense Enterprise**. Proprio come QlikView, Qlik Sense basa tutta la sua piattaforma su un motore in-memory che consente di mantenere le prestazioni che hanno reso famoso questo marchio. Le caratteristiche che contraddistinguono questo prodotto sono:

- *La visualizzazione interattiva.* Permette di creare report e dashboard in modo grafico trascinando direttamente all'interno dell'elaborato gli oggetti e i dati di interesse. A differenza degli altri prodotti di front end, Qlik prevede la possibilità di navigare nei dati attraverso semplici passaggi in modo da poter analizzarli secondo ogni prospettiva, ricercando nuove correlazioni e informazioni.
- *Interazione tra più sorgenti.* Qlik Sense si connette a più sorgenti dati, in modo da offrire una visualizzazione più completa. Questo prodotto è in grado di caricare i dati in tempo reale offrendo agli utenti dashboard e report sempre aggiornati.
- *Supporto mobile.* Qlik Sense supporta la visualizzazione e la modifica degli elaborati non solo sul desktop, ma anche su tablet e smartphone. In questo modo gli utenti possono essere sempre aggiornati sull'andamento delle aree aziendali di interesse.

Qlik Sense, oltre alla versione Enterprise, avente funzionalità più complete, offre altre due versioni, o meglio, due modalità in cui le aziende possono usufruire dei prodotti e delle funzionalità offerte da questa piattaforma:

- **Qlik Sense Cloud:** è la versione SaaS (*Software as a Service*) di Qlik Sense per poter usufruire di tutte le applicazioni e le potenzialità della piattaforma senza la necessità di un'installazione in locale. Gli utenti possono richiedere via web

solamente le funzionalità che servono a loro, senza dover necessariamente approcciarsi all'intera piattaforma. E' una soluzione che viene utilizzata dalle piccole imprese in quanto non consente una larga diffusione all'interno dell'azienda. Per quanto riguarda la versione in Cloud, l'offerta di Qlik è completamente gratuita, ma mette a disposizione uno spazio di archiviazione totale di soli 250 MB per un massimo di cinque utenti; rappresenta una soluzione molto limitata rispetto ad altri prodotti che supportano una versione in cloud.

- **Qlik Sense Desktop:** questa soluzione, a differenza di Qlik Sense Cloud, necessita di un'installazione, in quanto consiste in un'applicazione Windows che permette ai singoli utenti di creare, gestire e condividere report e dashboard interattivi.

Come già detto in precedenza, Qlik è leader indiscusso in ambito di data visualization e data discovering. L'aspetto sicuramente più significativo e positivo del prodotto sono i dashboard altamente interattivi. Secondo l'analisi effettuata da Gartner, Qlik è un prodotto leader nel magic quadrant, risultato messo in luce anche dalla soddisfazione dei clienti che vedono nella facilità di analisi una delle caratteristiche più importanti, oltre alla velocità e alla versatilità. Tra le note meno positive spicca la parte di reportistica, che nonostante lo sforzo e l'investimento fatto con Qlik Sense, resta ancora la parte più debole del prodotto, insieme all'analisi predittiva. A testimonianza di questo c'è l'impaginazione dei dashboard che è orientata alla visualizzazione e non alla stampa degli elaborati. Inoltre, offre una documentazione molto ampia e approfondita, anche se non è accompagnata da una sufficiente assistenza al cliente.

Riassumendo, Qlik è un ottimo prodotto per la creazione dei dashboard che, con la nuova piattaforma Qlik Sense, all'interno della quale sta cercando di integrare anche QlikView, sta investendo per coprire anche gli altri ambiti della BI. Nonostante questo impegno, è un prodotto che difficilmente può già essere considerato un prodotto di BI completo.

### 2.2.2. TIBCO Spotfire

Spotfire è una società di software informatici avente sede in Massachusetts. Nel 2007 questa società è stata acquisita e incorporata da TIBCO, una società californiana che produce soluzioni software in ambito della Business Management e Business Intelligence. L'acronimo TIBCO sta per "*The Information Bus Company*". Negli ultimi anni è stata una società molto attiva in ambito della BI. Nel 2014, infatti, ha acquistato

JasperSoft, un produttore software per la BI il cui omonimo prodotto verrà approfondito nel paragrafo successivo. Per quanto riguarda Spotfire, TIBCO offre tre soluzioni distinte: Spotfire Desktop, Spotfire Cloud e Spotfire Platform.

- **TIBCO Spotfire Desktop:** nel sito ufficiale del prodotto (<http://spotfire.tibco.com>) viene presentato come un software analitico concepito per l'elaborazione dei dati, che consente di estrapolare informazioni rilevanti e, quindi, un valore maggiore dai propri dati. E' una soluzione adatta ai singoli utenti che vede nella creazione e nella gestione dei dashboard il proprio punto di forza. Grazie a questo programma, gli utenti sono in grado di realizzare in maniera grafica e semplificata dei cruscotti analitici per la visualizzazione e l'esplorazione dei dati al fine di ricavare informazioni aggiuntive ed effettuare delle previsioni. L'interfaccia del programma è stata criticata da alcuni utenti in quanto visualizza più finestre di lavoro, risultando troppo piena durante lo sviluppo di un cruscotto. Nonostante questa piccola critica, Spotfire Desktop è un prodotto semplice ed intuitivo da utilizzare, grazie anche a delle strutture per la visualizzazione dei dati preimpostate come tabelle, tabelle a campi incrociati, grafici, diagrammi, mappe, caselle di testo e molto altro. Per poter effettuare l'analisi dei dati è necessario recuperarli dalle varie sorgenti che si hanno a disposizione, le quali possono essere semplici file o database. Spotfire supporta, inoltre, tecniche di analisi avanzate come la regressione lineare, l'estrapolazione e altro ancora.
- **TIBCO Spotfire Cloud:** è la versione in cloud di Spotfire Desktop; si tratta, infatti, di un software analitico as-a-service (SaaS), concepito per l'esplorazione dei dati. Oltre a tutte le funzionalità presenti anche nella versione Desktop, questo prodotto comprende la possibilità di eseguire visualizzazioni geo analitiche tramite delle mappe, al fine di offrire all'utilizzatore del cruscotto una rappresentazione dei dati ancora più chiara e intuitiva. Spotfire Cloud ha il vantaggio di mettere a disposizione tutto quello che è necessario per la realizzazione dei dashboard via web, senza dover installare nulla. Inoltre, rende molto più semplice la condivisione di tutti gli elaborati all'interno di una stessa realtà aziendale, centralizzando e certificando l'informazione. Spotfire offre fino a 250 GB di storage per il salvataggio dei dati su cui poter eseguire poi tutte le analisi.
- **TIBCO Spotfire Platform:** si tratta della versione più completa di questo prodotto, sviluppata per centralizzare tutte le informazioni nell'azienda e consentirne la condivisione. Offre una serie completa di strumenti, sia per la

visualizzazione dei dati, sia per l'analisi e l'esplorazione. Spotfire Platform comprende una gamma di funzionalità analitiche per l'analisi dei dati provenienti da singoli file, database esterni o dal cloud, che vengono integrati all'interno della piattaforma per ottenere informazioni più complete. Tra queste funzionalità le più significative sono: visualizzazione interattiva dei dati, analisi predittiva, indicatori KPI (Key Performance Indicator), analisi descrittiva, etc. Una volta realizzati i cruscotti, si possono condividere con le altre persone all'interno dell'azienda, non solo su PC, ma anche su dispositivi mobili come tablet e smartphone.

Il punto di forza di questo prodotto è senza dubbio la facilità sia di realizzazione dei cruscotti che di utilizzo di questi elaborati. E' un prodotto intuitivo che comprende numerose funzionalità di analisi e di previsione e che permette di navigare nei dati che si hanno a disposizione grazie a semplici passaggi. Se per prodotto completo si intende però un software che copre tutte le aree della BI, allora Spotfire non può essere considerato tale. Presenta infatti funzionalità ETL praticamente nulle, nonostante permetta di importare dati da diverse sorgenti, e anche la parte di reportistica non è a livelli di eccellenza. Tuttavia, questo non toglie il fatto che sia uno dei leader nella data visualization e nella data discovery, assieme agli altri prodotti che in questa analisi vengono definiti indipendenti. Altro aspetto da sottolineare è la poca integrazione di Spotfire con l'altro prodotto offerto da TIBCO, cioè JasperSoft. TIBCO ha acquisito numerosi prodotti che, messi insieme, ricoprono gran parte delle esigenze esistenti all'interno di un'azienda; il problema è che tutti questi prodotti non sono ancora amalgamati all'interno di una suite TIBCO. La società ci sta comunque lavorando e i presupposti per la realizzazione di questa piattaforma sembrano esserci. Ultima nota non propriamente positiva è la poca assistenza al cliente e lo scarso supporto all'utente.

### 2.2.3. Tableau

Tableau Software è una società di software americana fondata in California nel gennaio del 2003, trasferita successivamente a Seattle, Washington, dove è presente tutt'oggi la sede centrale. Lo scopo iniziale era quello di commercializzare uno studio effettuato dall'Università di Stanford sull'esplorazione grafica dei database relazionali. Quello che ne è risultato è una famiglia di prodotti per la visualizzazione grafica ed interattiva dei dati e per la navigazione delle basi dati sia relazionali che multidimensionali. Tableau è il principale prodotto di questa società, la quale, negli ultimi 4-5 anni, è andata incontro ad una crescita esponenziale sia in termini di diffusione in ambito aziendale sia, di

conseguenza, di fatturato. Questo notevole sviluppo è confermato anche dallo studio effettuato da Gartner sui prodotti per la BI. Dal 2014 Tableau è leader indiscusso del quadrante magico di Gartner, grazie ad una serie di caratteristiche che verranno approfondite in seguito. In generale, questo prodotto viene considerato leader, assieme a Qlik, della data visualization e della data discovery.

Tableau offre una serie di prodotti che hanno l'obiettivo di soddisfare nel miglior modo possibile le esigenze delle aziende. I prodotti sono: Tableau Desktop, Tableau Online, Tableau Server, Tableau Mobile, Tableau Public e Tableau Reader.

- **Tableau Desktop:** è un software per l'analisi dei dati; consiste in un ambiente che ricorda molto Excel, con cartelle e fogli di lavoro. Tra le caratteristiche più importanti di questo prodotto ci sono la facilità di apprendimento e di utilizzo e la possibilità di generare query senza dover scrivere del codice. Grazie ad una serie di best practices, già presenti all'interno del software, alla tecnologia drag-and-drop e ad una serie di tipologie di grafici a disposizione, l'utente è in grado di creare degli elaborati in modo grafico e intuitivo a partire da dati provenienti da diverse sorgenti, come database (IBM DB2, Microsoft Office Access, Microsoft SQL Server, etc.) e singoli file. Una volta che l'utente ha completato una serie di fogli di lavoro, ha la possibilità di integrarli in un unico cruscotto interattivo, all'interno del quale gli utilizzatori di questo elaborato potranno impostare filtri, selezionare solamente le aree di interesse ed esplorare i dati secondo le varie esigenze di analisi. E' inoltre possibile effettuare delle analisi matematiche, come ricavare le statistiche dai dati, calcolare tendenze, regressioni ed estrapolazioni. Tableau Desktop supporta anche le mappe interattive che offrono una visualizzazione dei dati suddivisi in aree geografiche, con la possibilità di selezionare e approfondire le informazioni di una particolare area. E' un prodotto per clienti che lavorano in locale e ha nella realizzazione dei cruscotti il suo punto di forza, come d'altronde anche tutti gli altri prodotti offerti da Tableau. Esistono due versioni di Tableau Desktop: *Personal* (con meno funzionalità e per singoli utenti) e *Professional* (con tutte le funzionalità e adatta per le aziende).
- **Tableau Server:** è un server basato sul web e sul mobile, facile sia per quanto riguarda l'installazione che l'utilizzo. L'utente può pubblicare i cruscotti realizzati attraverso Tableau Desktop in modo da avere tutte le informazioni condivise all'interno dell'impresa. Grazie a questo server, gli utenti possono aggiornare i propri dashboard sostanzialmente in tempo reale, oltre a poter apportare delle modifiche agli elaborati in tempi molto brevi. Non offre tutte le funzionalità che i

principali server, come Oracle e Microsoft, offrono, ma presenta tutte le caratteristiche di base necessarie per svolgere le operazioni richieste. In particolare, supporta la schedulazione degli elaborati, che permette di avere i dati sempre aggiornati all'istante di tempo determinato dagli utenti e offre un sistema di protezione tramite autorizzazioni per l'accesso attraverso un user name e una password, assegnati ad ogni utente di ogni livello.

- **Tableau Online:** è la versione in cloud di Tableau Server. E' un'applicazione SaaS (*Software as a Service*) che permette di centralizzare tutte le informazioni in un'unica architettura. Essa è sostanzialmente uguale a quella della versione Server e vi è possibile collegare tutti i dispositivi, anche mobili. I vantaggi di Tableau Online, rispetto a Tableau Server, sono quelli legati al cloud. Questo prodotto non necessita di un'installazione ed è altamente scalabile.
- **Tableau Mobile:** più che un prodotto si tratta di una funzionalità supportata sia da Tableau Server che da Tableau Online, grazie alla quale è possibile visualizzare, selezionare e filtrare i dati all'interno dei dashboard creati attraverso Tableau Desktop. Attraverso gli schermi touch dei tablet e degli smartphone, l'utente può effettuare della analisi e delle semplici operazioni sui dati come il drill down. Tableau Mobile è disponibile come un'applicazione per iPhone, iPad, Android e sui browsers dei dispositivi mobili.
- **Tableau Public:** è un servizio offerto da Tableau che permette di connettersi ai dati, creare visualizzazioni interattive e pubblicare l'elaborato direttamente nel proprio sito web. In questo modo, tutti i cruscotti condivisi saranno pubblici e quindi visibili a tutti.
- **Tableau Reader:** è un'applicazione desktop gratuita che permette all'utente di interagire con i cruscotti creati attraverso Tableau Desktop. Questa applicazione supporta operazioni sui dati quali visualizzazione, filtraggio, ordinamento, esportazione, impaginazione e stampa.

Come già detto in precedenza, Tableau è leader nell'ambito della parte front end della BI e, in particolare, della data visualization e della data discovery. E' un prodotto che ha visto negli ultimi anni una diffusione molto importante dovuta principalmente alla facilità di utilizzo, aspetto più significativo per gli utenti delle aziende, assieme ad una grafica che è la migliore offerta dal mercato. Grazie a tutte le componenti di questo prodotto, grandi quantità di dati provenienti da sorgenti differenti possono essere integrate mediante alcune funzionalità di back end e messe a disposizione degli utenti per la

creazione di visualizzazioni interattive, resa semplice da una serie di strutture già preimpostate. D'altra parte, ci sono degli aspetti che non sono all'altezza di un prodotto che viene considerato leader. Tableau presenta dei limiti per quanto riguarda la reportistica e le tipologie di analisi tradizionali. E' uno strumento fortemente orientato ai cruscotti e, di conseguenza, le altre funzionalità tipiche di uno strumento completo di BI sono leggermente carenti. Anche la gestione dei metadati non è tra le migliori che si possono trovare nel mercato. In conclusione, Tableau è senza dubbio leader nella data discovery, ma non si può ancora considerare un prodotto di BI completo.

### 2.3. Prodotti Open Source

Prima di effettuare l'analisi del settore del mercato di questa tipologia di prodotti, può essere utile chiarire alcuni concetti ad essi legati come *"open source"*, *"software libero"* e *"software gratis"*. Il primo concetto, che si è sviluppato a partire dagli anni '80, è quello di software libero. Con "libero" non si intende tanto la possibilità di utilizzare il programma in maniera indiscriminata (il software è comunque soggetto ad una licenza d'uso che ne regola l'utilizzo da parte degli utenti), ma consiste più in un'ideologia che in condizioni tecniche. E' un concetto che si basa sulla libertà dell'utente di eseguire il programma per qualsiasi scopo, di accedere al codice e modificarlo, di distribuirne copie al fine di aiutare il prossimo e di migliorarlo in modo che tutta la comunità ne tragga beneficio. A differenza del software libero, che si basa su principi come la libertà, l'aiuto e la comunità, il concetto di open source risulta molto più tecnico, cioè basato sulle caratteristiche della licenza. In questo caso l'utente ha la possibilità di accedere al codice sorgente potendo analizzarlo e, se lo ritiene necessario, modificarlo. Risulta evidente come i due concetti non siano diversi; si tratta sostanzialmente della stessa tipologia di prodotti visti con due prospettive differenti: da un lato vengono considerate le condizioni etiche e dall'altra quelle tecniche. Un programma, infatti, generalmente non può essere libero senza essere open source.

Il fatto che quasi sempre un prodotto open source sia anche gratis, porta molto spesso a considerare sinonimi i due termini, anche se in maniera scorretta. Questo è dovuto anche al fatto che, in ambito anglofono, il concetto di "libertà" è esprimibile con la parola *"free"*, che significa anche *"gratis"*.

Al giorno d'oggi si utilizza quasi esclusivamente il concetto di prodotto open source. Tra i tool selezionati in questo lavoro, alcuni offrono sia la versione Community, quindi gratis, sia versioni più complete aventi componenti e funzionalità aggiuntivi, disponibili attraverso un abbonamento e quindi aventi licenze non più libere, ma commerciali.

I prodotti selezionati sono quelli generalmente considerati i migliori per la loro tipologia e che presentano un peso importante anche a livello dell'intero mercato dei prodotti BI. Essi sono: Pentaho, TIBCO Jaspersoft BI e Eclipse BIRT Project.

### 2.3.1. Pentaho

Pentaho è una società relativamente giovane fondata nel 2004 da cinque soci con sede in Orlando, Stati Uniti d'America. Questa società offre una suite di BI Open Source in grado di coprire tutto l'ampio ventaglio di potenzialità della BI. All'interno del prodotto, si possono trovare funzionalità di reporting anche in modalità self-service, cubi OLAP per l'analisi multidimensionale del dato, cruscotti e dashboard per visualizzare in modo semplice e intuitivo i principali indicatori in grado di valutare l'andamento dell'azienda, integrazione dei dati attraverso funzionalità ETL per facilitare l'integrazione dei dati provenienti da sorgenti differenti, data mining.

La suite di Pentaho offre due prodotti: la Community Edition (CE) e l'Enterprise Edition (EE). La CE è la soluzione open source offerta da questa suite e, nonostante sia la versione gratuita, contiene al suo interno una serie di prodotti che offrono una valida alternativa ad altre soluzioni di BI. La EE fornisce delle componenti e dei programmi aggiuntivi che rendono il prodotto più potente e più competitivo anche per le imprese medio-grandi. La versione enterprise si ottiene attraverso un abbonamento annuale che comprende anche servizi di assistenza aggiuntivi. L'ammontare della licenza varia a seconda dei servizi richiesti dalle singole aziende, di conseguenza è difficile eseguire una stima.

In seguito verranno analizzati i principali prodotti che compongono le due tipologie di suite proposte da Pentaho. Si possono suddividere questi prodotti in due ulteriori tipologie: gli applications server e i desktop application.

I principali prodotti application server offerti da entrambe le suite sono:

- **Pentaho BI Platform:** viene più generalmente chiamato BI Platform e, recentemente, è stato anche rinominato in *Business Analytics Platform* (BA Platform). Include una serie di funzionalità che consistono nella gestione della sicurezza, esecuzione di report, visualizzazione di dashboard, script per la definizione di regole di business e analisi OLAP. Quest'applicazione viene eseguita in Apache Java Application Server.

- **Pentaho Analysis Services (Mondrian):** è un'application server scritto in java in grado di eseguire analisi OLAP (*Online Analytical Processing*). Supporta il linguaggio XML e MDX (*multidimensional expressions*) per le query ed è capace di leggere da sorgenti SQL e da altre sorgenti, aggregando i dati in una memoria cache.
- **Pentaho Data Access Wizard:** il plug-in è fornito assieme a tutti i server e permette all'utente di creare nuove sorgenti dati da utilizzare per tutto il sistema a partire da altri database o da file CSV presenti nel server. All'utente viene fornita la possibilità di creare un modello dei dati che descrive come i campi delle tabelle sono in relazione tra di loro. Questi modelli che possono essere creati dall'utente, sono messi a disposizione di altri prodotti che ne permettono l'interrogazione tramite query.

Oltre a questi servizi offerti dalla Community Edition, sono resi disponibili agli abbonati della versione Enterprise altre applicazioni server:

- **Pentaho Dashboard Designer (PDD):** plug-in che consente agli utenti di creare dashboard che forniscono una visione centralizzata degli indicatori chiave di performance e altri aspetti del business. E' supportata un'interfaccia che permette agli utenti di creare dashboard in modo grafico, trascinando i vari oggetti di interesse.
- **Pentaho Analyzer:** fornisce una piattaforma web grazie alla quale l'utente può creare graficamente query MDX ed esportare la tabella risultante in formato pdf o xls. E' noto per poter lavorare su Apple attraverso il browser Safari.
- **Pentaho Interactive Reporting (PIR):** permette di creare report ad hoc in modo grafico trascinando gli oggetti di interesse.
- **Pentaho Mobile:** novità introdotta recentemente che permette di usufruire delle principali funzionalità di analisi OLAP e gestione report e dashboard attraverso piccoli schermi touch screen.

Oltre agli application server, Pentaho offre una serie di applicazioni desktop per gli utenti, tutte disponibili sia nella CE che nella EE:

- **Pentaho Data Integrator (PDI):** generalmente chiamato Kettle, consiste in un motore ETL che permette l'estrazione, l'integrazione e il caricamento di dati provenienti da sorgenti differenti attraverso un'interfaccia grafica semplificata.

Essa permette all'utente di bypassare il puro codice e utilizzare oggetti grafici per rappresentare il flusso di lavoro.

- **Pentaho for Big Data:** è un tool ETL che si basa su Kettle; permette di eseguire job su grandi quantità di dati provenienti da sorgenti come Amazon e altre fonti di dati NoSQL.
- **Pentaho Report Designer:** è un prodotto che permette la generazione di report attraverso interrogazioni di sorgenti differenti. Anche in questo caso, l'utente è facilitato dalla presenza di un'interfaccia grafica semplificata.
- **Pentaho Data Mining:** permette l'elaborazione dei dati, analisi di regressione lineare, metodi di classificazione, interpolazioni... Sulla base dei modelli individuati dall'attività di data mining, gli utenti possono eseguire una previsione degli eventi futuri.
- **Pentaho Metadata Editor (PME):** usato per creare modelli di business; rappresenta un livello di astrazione dell'architettura fisica delle sorgenti. E' uno strumento molto importante a livello aziendale e la sua presenza in una soluzione open source è sicuramente un aspetto positivo.

Pentaho è un prodotto già molto utilizzato nelle varie realtà aziendali ed è in continuo sviluppo. Presenta delle caratteristiche molto importanti soprattutto per le PMI e può rappresentare un'ottima alternativa a delle soluzioni offerte da prodotti magari molto più costosi. Tra gli aspetti positivi che spingono le aziende ad orientarsi verso questo prodotto (individuati anche da Gartner nel documento pubblicato con l'analisi del mercato dei software BI del 2015), c'è sicuramente il basso costo delle licenze e, di conseguenza, un elevato rapporto qualità-prezzo. Il prodotto che spicca maggiormente, tra tutti quelli offerti, è Kettle. Questo strumento ETL, considerando il fatto che non richiede una licenza, offre un'ottima alternativa alle soluzioni proposte dai prodotti storici, anche se le potenzialità sono ovviamente inferiori. Altri aspetti positivi sono la messa a disposizione agli utenti di una soluzione in cloud, scalabile e flessibile, che permette alle aziende di non dover investire in un'infrastruttura hardware costosa, e la semplicità dei tool, grazie all'utilizzo di interfacce grafiche che rendono gli utenti sempre più indipendenti e in grado di gestire l'infrastruttura senza l'aiuto di tecnici specializzati. Ci sono, d'altro canto, anche degli aspetti negativi come, ad esempio, la poca affidabilità del prodotto, messa in luce sempre nello studio di Gartner, secondo il quale il 24% dei clienti lo ritiene inaffidabile e il 68% dei clienti ha riscontrato dei problemi di vario genere con la piattaforma. Altro problema riscontrato dai clienti è la scarsa assistenza fornita alle aziende.

### 2.3.2. TIBCO Jaspersoft BI

Jaspersoft BI è una suite di Business Intelligence caratterizzata da funzionalità complete, architettura leggera e flessibile e dai costi contenuti, se non nulli. Jaspersoft è un progetto nato nel 2001 e portato avanti grazie a finanziamenti provenienti da importanti società. Nel 2014, la TIBCO (società californiana che produce soluzioni software per le aree di Business Management e Business Intelligence) ha acquistato Jaspersoft per una cifra pari a circa \$185 milioni.

Come descritto nel sito web ufficiale (<https://www.jaspersoft.com>), Jaspersoft fornisce servizi di reporting, dashboard, analisi e integrazione dei dati, adatti a qualsiasi soluzione di business. Un aspetto importante che caratterizza questo prodotto e che permette alle aziende di migliorare i processi decisionali, è l'utilizzo di un'interfaccia utente web-based che facilita l'uso degli strumenti di analisi messi a disposizione per l'utente. Jaspersoft BI suite include i seguenti prodotti:

- **JasperReports Library:** è uno dei motori di reporting open source più popolari, è scritto interamente in Java ed è in grado di utilizzare dati provenienti da qualsiasi sorgente per creare documenti esportabili o stampabili in più formati, quali HTML, PDF, Excel, OpenOffice e Word.
- **Jaspersoft iReport Designer:** tool in grado di creare layout molto sofisticati contenenti grafici, immagini, sottoreport, testi, tabelle a campi incrociati, etc. E' possibile accedere ai vari database attraverso diverse forme di connessione e pubblicare i report in molti formati differenti come PDF, XML, XLS, CSV, HTML, TXT, e DOCS. A partire dall'inizio del 2016, questo componente è stato sostituito da **Jaspersoft Studio** il quale è comunque in grado di leggere e modificare i report creati con *iReport Designer*.
- **JasperReports Server:** è un server autonomo e integrabile; fornisce reporting e analisi che possono essere integrati in un'applicazione web o mobile. Questo componente è ottimizzato per condividere, proteggere e gestire i report e le viste eseguite con questa piattaforma. Comprende, inoltre, una serie di strumenti per la creazione e la visualizzazione di dashboard altamente interattivi, grazie ai quali è possibile rappresentare i dati in maniera grafica. E' anche possibile navigare all'interno delle informazioni rappresentate, al fine di analizzare le varie aree aziendali di interesse.
- **Jaspersoft ETL:** costituisce il tool back-end della piattaforma e consente l'estrazione, l'integrazione e il caricamento dei dati da sorgenti differenti in un DW

o in uno specifico Data Mart. Permette di definire i flussi di lavoro e può collegarsi sia a sorgenti proprietarie che a quelle aperte.

- **Jaspersoft OLAP:** permette di manipolare, modellare e visualizzare qualsiasi tipo di dato attraverso un'analisi di tipo multidimensionale (OLAP), al fine di individuare i problemi e i trend nell'azienda e aiutare i decision maker a prendere migliori decisioni in tempi più brevi.

La piattaforma BI di Jaspersoft è una piattaforma completa dal punto di vista delle funzionalità, che vanno dall'estrazione dei dati dalle sorgenti fino all'analisi dei Big Data e le analisi multidimensionali.

Esistono cinque versioni della suite di Jaspersoft: Community, Reporting, AWS, Professional ed Enterprise. Per le finalità di quest'analisi, vale la pena soffermarsi solamente sulla versione Community e su quella Enterprise. La prima, che comprende le funzionalità di base della piattaforma Jaspersoft, è soggetta ad una licenza AGPL e rientra quindi nella tipologia di software open source, mentre la seconda è quella più completa che, però, prevede una licenza commerciale basata sul numero di core CPU forniti all'utente. Chiaramente, questa versione offre molte funzionalità aggiuntive, come il supporto tecnico professionale, reporting ad hoc, visualizzazioni basate su Flash e molto altro.

La suite BI Jaspersoft è sicuramente una piattaforma orientata al front-end e, quindi, al reporting e all'analisi dei dati, ma presenta al suo interno una serie di tool che la rendono completa in ambito BI, come lo strumento ETL per la creazione di Data Warehouse e Data Mart e la gestione dei metadati. Non mancano, però, alcuni aspetti negativi, messi in evidenza anche dallo studio effettuato da Gartner del mercato degli strumenti BI: la complessità di apprendimento da parte degli utenti, nonostante la presenza di interfacce grafiche, e l'assistenza e il supporto al cliente che non sono all'altezza di altri software concorrenti nel mercato.

### 2.3.3. Eclipse BIRT Project

Nel sito ufficiale (<http://www.eclipse.org/birt>) si trova la definizione di BIRT (Business Intelligence and Reporting Tools) Project, cioè un progetto open source fornito dalla piattaforma tecnologica BIRT per creare visualizzazioni di dati e report, che possono essere incorporate in ricche applicazioni client e web. E' un progetto software di alto livello all'interno della Eclipse Foundation, un consorzio indipendente e no-profit di fornitori del settore software e una comunità open source. Il progetto è sponsorizzato da

Actuate, assieme al contributo di IBM e Innovent Solutions ed è rilasciato sotto la *Eclipse Public License* (EPL).

Il principale obiettivo di questo progetto è quello di affrontare una vasta gamma di esigenze di reporting che vanno dal mondo operativo fino alle analisi multidimensionali (OLAP). Inizialmente il progetto, approvato nel 2004, si era concentrato sulle funzionalità che consentivano agli sviluppatori di applicazioni di progettare e integrare report all'interno delle applicazioni stesse, quando, invece, gli utenti non avevano generalmente a disposizione le conoscenze per sviluppare con un linguaggio di programmazione. Successivamente, già con la seconda versione, il prodotto è cresciuto molto, mettendo a disposizione nuovi report, un numero maggiore di grafici e la possibilità di creare dei modelli da poter riutilizzare in seguito.

BIRT comprende due componenti principali: un report designer basato su Eclipse e un componente runtime per la generazione di report implementabile in qualsiasi ambiente Java. Include, inoltre, un motore per produrre e integrare grafici all'interno delle applicazioni. Le varie tipologie di visualizzazione dei dati che è possibile inserire nelle applicazioni sono: liste di dati, grafici (rappresentazioni grafiche dei dati), tabelle a campi incrociati (visualizza i dati in due dimensioni), lettere e documenti. Spesso un report potrebbe necessitare di più tipologie tra quelle sopracitate; in questi casi i report vengono definiti composti.

I report sono costituiti da quattro parti principali:

- *Dati*: Database, servizi web, classi Java, sono tutte possibili fonti dati per i report BIRT. Inoltre, un singolo report è in grado di includere dati provenienti da più sorgenti differenti, che possono essere messe in relazioni tra di loro grazie all'utilizzo di join.
- *Trasformazione dati*: I report devono presentare i dati aggregati, filtrati, ordinati e riepilogati in base alle esigenze degli utenti. BIRT permette operazioni sofisticate, quali: raggruppamento su somme, percentuali di totali complessivi e molto altro su dati provenienti da sorgenti diverse.
- *Logiche di business*: I dati a disposizione di una azienda non sono generalmente già strutturati secondo le proprie esigenze. Molti report richiedono una logica di business specifica, in modo tale da trasformare dati grezzi in informazioni utili agli utenti. Se la logica è propria del singolo report, è possibile implementarla al suo interno grazie al supporto offerto da BIRT per programmare in Java. Se,

invece, l'applicazione contiene già la logica, è possibile richiamarla nel codice Java esistente.

- **Presentazione:** una volta che i dati sono stati ripuliti e organizzati, il tool offre una vasta gamma di opzioni per la presentazione dei dati come tabelle, grafici, testo... La visualizzazione grafica dei dati offre una modalità più semplice e veloce di comprensione del tema in esame.

Il progetto BIRT comprende molti componenti e molti plug-in che vanno a completare le funzionalità delle soluzioni offerte. Di seguito verranno presentati i componenti:

- **BIRT Report Designer:** è un componente di Eclipse usato per la progettazione di report salvati in un formato XML. Questo componente può essere scaricato come una piattaforma rich client (RCP), cioè uno strumento di programmazione che rende più facile l'integrazione di più componenti software indipendenti. Il trattamento dei dati avviene principalmente sul lato client, oppure come un all in one download di Eclipse.
- **Design Engine:** è il motore responsabile della creazione e della modifica dei report. L'API (*Application Programming Interface*) del Design Engine viene utilizzata da questo componente per la creazione delle presentazioni XLM.
- **Report Engine:** svolge la funzionalità di generare i report. Utilizzando l'API del Report Engine, il motore può essere integrato all'interno di qualsiasi applicazione Java.
- **Charting Engine:** è utilizzato per generare grafici sia in stand-alone che all'interno di report BIRT. Grazie all'API del Charting Engine, il Design e il Report Engine permette di fornire grafici di vario tipo.
- **BIRT Viewer:** il progetto BIRT fornisce questo componente che permette di visualizzare i report all'interno di Eclipse. L'output fornito da questo viewer può essere in formato HTML, PDF, XLS, DOC, PPT. Inoltre, l'utente può esportare i dati del report in formato CSV e stamparli.

BIRT è, attualmente, una delle piattaforme più usate per la visualizzazione e il reporting dei dati. Inoltre, può contare su un gruppo di sviluppatori che sta crescendo sempre di più e che rappresenta molte delle più importanti organizzazioni come, per esempio, IBM, anche se negli ultimi anni lo sviluppo di questa piattaforma ha subito un leggero rallentamento rispetto alle due concorrenti open source trattate in questo confronto. Tra gli aspetti positivi di BIRT, c'è sicuramente la capacità di integrarsi con molte sorgenti

dati in diversi ambienti. In particolar modo, è in grado di integrare dati provenienti da database SQL, che sono tra i più diffusi, ma anche da altri tipi di sorgenti. BIRD Report Designer ha a sua disposizione anche un ricco set di funzionalità che, grazie all'interfaccia utente intuitiva, rende questo programma uno dei più usabili, nonostante inizialmente fosse orientato principalmente ai programmatori. Una differenza importante con gli altri strumenti sta nella visualizzazione dei report via web.

Al termine di questa panoramica, viene proposta una tabella riassuntiva dei prodotti offerti dalle soluzioni analizzate in questo capitolo, in modo da avere una visione più chiara di tutto l'insieme di opportunità che il mercato mette a disposizione.

PRODUTTORE	ETL e DW	Reporting	Dashboarding	Predictive analysis
ORACLE	Oracle Data Integrator	Oracle Hyperion Oracle Business Intelligence	Oracle Hyperion Oracle Business Intelligence	Oracle Hyperion
MICROSOFT	SQL Server Integration Services	Microsoft BI Semantic Model SQL Server Reporting Services Microsoft Power BI	Microsoft Power BI	Microsoft Power BI
IBM	IBM InfoSphere DataStage	IBM Cognos	IBM Cognos	IBM SPSS
SAP	SAP Data Services SAP Business Warehouse	SAP Business Objects BI SAP Crystal Report SAP Business Warehouse	SAP Crystal Dashboard	-
MICROSTRATEGY	MicroStrategy Intelligence Server	MicroStrategy Architect MicroStrategy Desktop MicroStrategy Web	MicroStrategy Desktop MicroStrategy Web	MicroStrategy Desktop MicroStrategy Web
QLIK	Qlik Sense (ETL) Qlik View (ETL)	Qlik Sense Qlik View	Qlik Sense Qlik View	-
SPOTFIRE	-	TIBCO Spotfire Desktop TIBCO Spotfire Cloud TIBCO Spotfire Platform	TIBCO Spotfire Desktop TIBCO Spotfire Cloud TIBCO Spotfire Platform	TIBCO Spotfire Platform
TABLEAU	-	Tableau Desktop Tableau Server	Tableau Desktop Tableau Server	-
PENTAHO	Pentaho Data Integrator	Pentaho Interactive Reporting Pentaho Report Designer	Pentaho Dashboard Designer	Pentaho Data Mining
JASPERSOFT	Jaspersoft ETL	JasperReports Library Jaspersoft iReport Designer JasperReports Server	JasperReports Server	-
BIRT	-	Eclips BIRT Project	Eclips BIRT Project	-

Tabella 2.2: Prodotti analizzati suddivisi in base all'area della BI

# CAPITOLO 3

## Comparazione software

### 3.1. PMI: Definizioni

Le piccole e medie imprese (PMI) sono aziende aventi dimensioni che, in termini occupazionali e finanziari, rientrano all'interno di certe soglie prefissate dall'Unione Europea. Fino alla seconda metà degli anni '90, ogni stato applicava la propria definizione di PMI, il che rendeva difficile il coordinamento tra le iniziative dell'UE a favore di questa tipologia di imprese. Con la Raccomandazione 2003/361/CE del 6 maggio 2003, la Commissione Europea ha uniformato i limiti per la classificazione delle aziende. Questo documento (che si può trovare nel sito: <http://eur-lex.europa.eu/>) afferma: "Una Media impresa è definita come un'impresa il cui organico sia inferiore a 250 persone e il cui fatturato non superi 50 milioni di euro o il cui bilancio annuale non sia superiore a 43 milioni di euro. Una piccola impresa è definita come un'impresa il cui organico sia inferiore a 50 persone e il cui fatturato o il totale del bilancio annuale non superi 10 milioni di euro. Una microimpresa è definita come un'impresa il cui organico sia inferiore a 10 persone e il cui fatturato o il totale del bilancio annuale non superi 2 milioni di euro."

Tipo	Occupati		Fatturato (Milioni di €)		Totale di bilancio (Milioni di €)
Media impresa	< 250	e	≤ 50	oppure	≤ 43
Piccola impresa	< 50	e	≤ 10	oppure	≤ 10
Micro impresa	< 10	e	≤ 2	oppure	≤ 2

Tabella 3.1: riassunto definizione imprese (Fonte: <https://it.wikipedia.org>)

La diffusione di queste realtà dipende da stato a stato e l'Italia è caratterizzata da una grande diffusione delle PMI, aspetto spesso associato ad una debolezza del mercato. Queste aziende, infatti, sono molto influenzate dalle grazie imprese, dalle quali vengono spesso sovrastate ed è per questo motivo che necessitano di finanziamenti esterni per poter rimanere competitive all'interno del mercato. Risulta ancor più evidente, quindi, come la scelta di un'infrastruttura informatica adeguata possa fare la differenza tra crescita e fallimento.

### 3.2. Tecnologie database

Fin dalla loro introduzione negli anni '80, i Relational Database Management Systems (RDBMS) sono diventati i database management systems standard in ambito aziendale. La scelta della base dati da utilizzare e su cui, di conseguenza, basare tutta l'infrastruttura informatica dell'azienda, è una decisione a lungo termine che l'IT manager deve prendere. In base a questa scelta dipenderanno tutte le ulteriori decisioni di medio e breve termine, come la scelta dei vari prodotti da utilizzare sia in fase di back end che in quella di front end. E' una decisione molto delicata che deve essere raggiunta al termine di un processo ben preciso costituito da più fasi:

- *Analisi*: in questa fase l'IT manager deve analizzare le esigenze aziendali per determinare le reali problematiche e i temi da affrontare. Devono essere individuati dei fattori che saranno fondamentali per la scelta, come il volume dei dati da trattare, il numero di utenti che dovrà accedere al database, il tipo di dati e, di conseguenza, la loro modellazione logica. Inoltre, devono essere individuati i fatti di interesse per l'azienda che dovranno essere poi rappresentati all'interno della base dati.
- *Comparazione*: una volta definite le necessità e le specifiche tecniche di cui il database necessita, l'IT manager, grazie ad una conoscenza delle soluzioni offerte dal mercato, effettua una comparazione tra i vari prodotti al fine di individuare quello che meglio si adatta alle esigenze, tenendo conto dei vincoli aziendali che ci possono essere, come ad esempio i costi.
- *Scelta*: al termine dell'analisi e della comparazione, si deve effettuare una scelta che, a questo punto, sarà ponderata e ben pensata.

La prima fase dipende dalla singola azienda in quanto, come già detto anche in precedenza, le esigenze dipendono dalle singole realtà e anche le regole interne e di

business sono soggettive. Quello che può essere comune a tutte le aziende è lo studio del mercato e il confronto tra le varie soluzioni offerte.

Un aiuto molto importante per una prima analisi dei RDBMS (ma anche dei vari tool di business intelligence che verranno confrontati nel paragrafo successivo) viene fornito dallo studio che ogni anno viene effettuato da Gartner, una società multinazionale leader mondiale nella consulenza e fondata nel 1979. Il risultato di queste analisi, che appunto possono riguardare il mercato di software e di prodotti di diversi settori, comprende un report in cui vengono approfondite le caratteristiche dei vari vendor e un grafico definito *Gartner Magic Quadrant* (Quadrante Magico di Gartner). In questo quadrante, le aziende che offrono i vari prodotti vengono posizionate in base a due criteri, che corrispondono rispettivamente all'asse delle ascisse e a quello delle ordinate: "completezza di visione" e "capacità di esecuzione". La "completezza di visione" comprende otto criteri di valutazione: comprensione del mercato, strategia marketing, strategia vendite, strategia nell'offerta, modello di business, strategia di mercato, innovazione e strategia geografica. Il criterio "Capacità di esecuzione" include, invece, fattori come la dinamicità dell'azienda, le operations, l'esecuzione nelle vendite, la politica dei prezzi e l'esecuzione in termini di marketing. I vendor vengono quindi posizionate nel grafico in base al punteggio che ottengono in questi due criteri. Gartner individua di conseguenza quattro categorie di aziende:

- *Leaders*: sono quelle che hanno ottenuto i punteggi più alti in entrambi i criteri, posizionandosi quindi nel primo quadrante di quello che risulta essere un piano cartesiano a tutti gli effetti. I *leaders* offrono tipicamente i prodotti di maggior spessore nel mercato aventi l'offerta maggiore e i fatturati più importanti.
- *Challengers*: sono caratterizzate da un punteggio positivo per quanto riguarda la capacità di esecuzione e da uno negativo nella completezza di visione.
- *Visionaries*: sono nella situazione opposta ai challengers, cioè con un punteggio positivo nella completezza di visione e negativo nella capacità di esecuzione. Sono principalmente aziende di piccole dimensioni.
- *Niche players*: aziende caratterizzate da un punteggio basso in entrambi i criteri. Generalmente corrispondono alle realtà che sono entrate da poco all'interno del mercato in esame.

Sebbene lo studio di Gartner rappresenti un aiuto molto importante in ambito decisionale per le aziende, esso non deve essere l'unico strumento utilizzato, ma il punto di partenza. Risulta evidente che un prodotto non può essere scelto e classificato in base a due soli

criteri; per effettuare una scelta basata su un'analisi approfondita è necessario valutare tutta una serie di criteri che vanno ad individuare come un prodotto risponde ai requisiti imposti in ambito aziendale. Senza addentrarsi nelle singole specifiche richieste da un'azienda, è possibile individuare una serie di parametri di valutazione in base ai quali è possibile effettuare la scelta della tecnologia database. Tra questi criteri, quelli che, generalmente, hanno il maggior peso per le piccole e medie imprese sono:

- *Architettura del database*, sia fisica (celle di memoria, ...) che logica (tabelle, ...);
- *Facilità di utilizzo*, uno degli aspetti più importanti e determinanti nella scelta sia del database che in generale di tutti i software;
- *Scalabilità*, intesa come la possibilità, da parte dell'utente, di poter integrare la soluzione inizialmente adottata con funzionalità aggiuntive in base alle necessità che man mano possono nascere all'interno di un progetto;
- *Licenze*, legate in modo diretto al parametro precedente; altro aspetto fondamentale per la scelta del DB;
- *Compatibilità con i vari tool di BI*, non solo con quelli appartenenti allo stesso vendor.

Ai fini di questo lavoro, verranno confrontati due RDBMS che sono già stati citati nel capitolo precedente durante la panoramica dei prodotti storici: Oracle Database e Microsoft SQL Server. Il mercato, ovviamente, offre un numero molto grande di alternative a questi due prodotti, che vanno da soluzioni di altre aziende molto importanti (come, per esempio, DB2, il database di IBM) a soluzioni open source (ad esempio MySQL di Oracle). Si stanno inoltre diffondendo sempre più delle tecnologie NoSQL che consistono in database non relazionali, ma che si basano sulle teorie dei grafi, come, ad esempio, Neo4j e MongoDB. Nonostante tutte queste alternative, si è scelto di prendere in considerazione solo gli RDBMS di Oracle e Microsoft, in quanto sono le due soluzioni maggiormente impiegate in ambito aziendale, caratterizzate da una presenza costante nel mercato da ormai molti anni e da fatturati che sono tra i più alti in questo ambito.

L'offerta di Oracle è tra le più complete e vaste in assoluto; la società, infatti, ha da sempre adottato la filosofia di cercare di soddisfare tutte le tipologie di esigenze presenti all'interno di un'impresa, pensando soluzioni per le grandi aziende come per quelle più piccole. La versione più recente di Oracle, inteso come database tradizionale, è **Oracle Database 11g**, che è attualmente una delle versioni più diffusa all'interno delle aziende. L'ultima novità introdotta da Oracle Corporation, invece, è un database in cloud che svolge la funzione di *Database Container*. Nelle versioni tradizionali, la memoria e i

processi sono allocati per ciascun database in server locali; nella versione in cloud, invece, si è concentrata tutta la parte di memorizzazione all'interno di un unico database che svolge la funzione di un vero e proprio container, abbassando in modo significativo le infrastrutture hardware e i costi operativi. Su questo nuovo principio, Oracle ha lanciato nel mercato nel giugno del 2013 **Oracle Database 12c** (dove la c sta per cloud).

Un server Oracle è rappresentato da due strutture: il database (DB) e l'istanza. Con il termine database si indicano i file fisici in cui sono memorizzati i dati, mentre per istanza si intende l'insieme delle aree di memoria e dei processi necessari ad accedere ai dati memorizzati nel database. Il DB consiste in strutture logiche di memorizzazione dei dati (tabelle, indici, etc.) e di strutture fisiche che contengono quelle logiche. Per quanto riguarda l'architettura logica, si ha che ogni database è suddiviso in uno o più *Tablespace*, cioè strutture che permettono di suddividere tutti i dati in base alle esigenze e le aree di interesse. Questo permette di gestire singole aree del DB senza dover intervenire per forza su tutto il sistema. Ogni *Tablespace* è costituito da più *Segment* che corrispondono alle tabelle e agli indici memorizzati nel database. A loro volta i *Segment* sono formati da uno o più *Extent*, cioè blocchi contigui di *Data Block* che sono la più piccola unità di memorizzazione in Oracle. La dimensione dei block viene decisa dall'amministratore del database nel momento della creazione di quest'ultimo. Generalmente, viene utilizzata una dimensione di 8 Kb e la dimensione più piccola supportata da Oracle è di 2 Kb. Per quanto riguarda l'architettura fisica, invece, ogni database è suddiviso in uno o più *Data file* che contengono tutti i dati del DB. Ciascun *data file* è associato esclusivamente ad un solo *tablespace*, ma un *tablespace* può essere formato anche da più di un *data file*. L'archiviazione fisica dei file avviene in blocchi di celle di memoria in cui vengono salvati i file contenuti dei *data block*, ai quali possono corrispondere più celle fisiche.

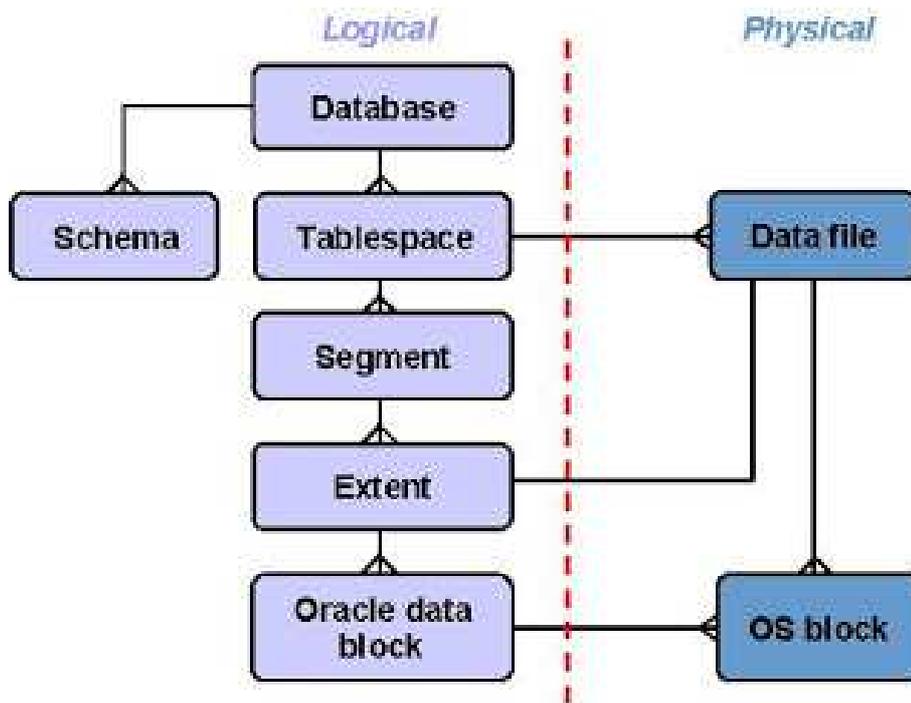


Figura 3.1: Architettura logica e fisica del database Oracle. (Fonte: [oracle-sliit.blogspot.com](http://oracle-sliit.blogspot.com))

Per quanto riguarda Microsoft, invece, l'ultimo prodotto lanciato nel mercato è Microsoft SQL Server 2014, un DBMS relazionale, sviluppato in ottica dei big data e del cloud. Per quanto riguarda la struttura fisica del database stesso e dei suoi oggetti, Microsoft SQL Server risulta sostanzialmente simile ad Oracle. Questo server si suddivide in due tipologie di database contenuti al suo interno: i DB di sistema, all'interno dei quali vengono memorizzate tutte le informazioni e gli oggetti necessari al funzionamento di SQL Server, e i DB utente, dove vengono salvati tutti i dati appartenenti agli utenti e gli oggetti del database che sono simili a quelli presenti in Oracle (tabelle, indici, viste, data type, procedure memorizzate, trigger, etc.). Per ognuno di questi elementi, vengono memorizzate nei metadati le informazioni riguardanti la creazione, le modifiche e le eliminazioni.

Microsoft ha rilasciato 6 versioni di SQL Server 2014:

- Express: è la versione entry-level gratuita, consigliata a chi deve gestire server di piccole dimensioni, adatto allo sviluppo di piccole applicazioni desktop.
- Developer: consente agli sviluppatori di creare qualsiasi tipo di applicazione basata su SQL Server; può però essere utilizzato come sistema di sviluppo e test e non come server di produzione.

- **Web:** è una soluzione basata su web che fornisce all'utente scalabilità, affidabilità e la possibilità di gestione della propria struttura.
- **Standard:** fornisce soluzioni base per il data management e la BI attraverso database in grado di supportare le applicazioni per le piccole aziende.
- **Business Intelligence:** soluzione che verrà presa in esame in seguito in quanto più idonea alle PMI. Offre una piattaforma completa che permette alle organizzazioni di creare e distribuire soluzioni di BI sicure e scalabili.
- **Enterprise:** versione top di gamma di Microsoft SQL Server pensata e sviluppata per le grandi imprese che necessitano di gestire enormi quantità di dati e sviluppare soluzioni in grado di ricoprire tutte le esigenze delle varie aree di interesse aziendale.

Di per sé, le due tecnologie sono sostanzialmente simili dal punto di vista dell'architettura; gli oggetti del database, infatti, sono gli stessi e anche la loro rappresentazione logica, con qualche piccola differenza, tutto sommato coincide. Tuttavia, si possono individuare tre principali differenze tecniche nei due prodotti:

- Il linguaggio attraverso il quale è possibile interagire con il database e interrogarlo. Oracle utilizza una variante del linguaggio SQL definita PL-SQL (*Procedural Language SQL*), mentre Microsoft utilizza un'altra variante dello stesso linguaggio chiamata T-SQL (*Transact SQL*). Le principali differenze tra i due linguaggi sono la gestione delle variabili e la complessità, maggiore per il PL-SQL.
- Gestione delle transazioni. Su SQL Server, nel momento in cui la transazione finisce con il *commit*, le modifiche effettuate sono salvate e non è più possibile effettuare il *roll back*. In Oracle, invece, ogni connessione al DB viene considerata una transazione, per cui è possibile effettuare delle modifiche senza dover per forza fare il *commit*, che invece può essere fatto alla fine della transazione.
- Suddivisione logica degli oggetti del database. Su Microsoft gli oggetti del database appartengono al database al quale l'utente si collega; in Oracle, invece, ad ogni utente corrisponde uno schema al quale appartengono gli oggetti. In questo caso, l'utente si collega allo schema e non all'intero database.

Tra gli aspetti più importanti di Oracle, messi in luce anche dall'analisi di Gartner, c'è l'ampiezza della gamma delle soluzioni. Oracle offre soluzioni a qualsiasi tipo di esigenza, accompagnate dalle funzionalità più potenti che si possano trovare nel

mercato. Questo, però, va a discapito della facilità di utilizzo: tanto sono vaste le potenzialità di questo prodotto, quanto è difficoltoso per degli utenti imparare ad utilizzarlo in maniera autonoma per poterle sfruttare al meglio. D'altro canto, Microsoft fa della facilità di apprendimento e di utilizzo il suo punto di forza. Grazie ad una serie di interfacce grafiche semplici ed intuitive, Microsoft SQL Server offre prestazioni leggermente inferiori ad Oracle, ma comunque di ottimo livello. Gli utenti, generalmente, sono attirati da un lato dalle prestazioni di Oracle e dall'altro dalla semplicità e dalla fascia di prezzo leggermente inferiore di Microsoft. Si tratta delle due soluzioni più utilizzate alle quali, infatti, corrispondono i due fatturati maggiori. La decisione tra queste due alternative dipende quasi esclusivamente dalle scelte aziendali, sia in termini di costi che di prodotti da utilizzare in ambito gestionale e di business intelligence.

### **3.3. Parametri Software Selection**

Il procedimento per effettuare una scelta ponderata dei software di BI da appoggiare sull'infrastruttura del DB è del tutto analogo a quello per la scelta del RDBMS. La prima cosa da analizzare sono le esigenze dell'azienda, dalle quali è possibile estrarre tutti i parametri che serviranno per confrontare i vari prodotti. Solo se l'azienda riuscirà a comprendere le reali necessità e a dare loro la giusta importanza, potrà scegliere i software adeguati. Risulta evidente come uno degli aspetti che subito viene preso in considerazione sono le licenze, cioè i costi di acquisto del prodotto. Il budget che un'azienda decide di dedicare a tutta l'infrastruttura informatica è uno dei parametri più significativi; sempre più frequentemente, infatti, molte aziende, per ridurre i costi, adottano prodotti che non corrispondono alle esigenze pratiche, andando a discapito di tutta la gestione e l'analisi dei dati. Vista la grande complessità delle licenze che caratterizza i prodotti e la possibilità di ricevere dei preventivi solo in presenza di un progetto concreto da parte di un'azienda, quest'analisi non considererà in modo approfondito questo parametro. Sicuramente verranno fatte delle considerazioni generali sui prezzi, tenendo presente la grande variabilità che li contraddistingue.

#### **3.3.1. Parametri**

Ai fini di questo lavoro, sono stati individuati dieci criteri in base ai quali verranno giudicati i vari prodotti analizzati nel capitolo precedente. Si è cercato di comprendere in questi parametri tutte le varie aree di interesse per una PMI, che vanno dall'architettura del

programma fino al supporto dei vari ambiti della BI, come la parte di back end e tutta la parte di front end.

I dieci parametri presi in considerazione sono:

**1. Infrastruttura e architettura.**

Con questa voce si intende la completezza del prodotto, cioè il supporto di tutte le funzionalità che ci si aspetta da uno strumento di BI. Un prodotto che non prevede nessuna funzionalità di integrazione può, sicuramente, essere considerato un software leader per esempio nella data visualization, ma non può essere considerato uno strumento di BI completo così definito. Inoltre, viene presa in considerazione anche da quali sistemi operativi è supportato.

**2. Connettività.**

Valutazione della capacità del prodotto di connettersi, e quindi di estrarre dati, da sorgenti differenti. E' un aspetto molto importante, che permette alle aziende di avere a disposizione dati differenti resi omogenei e disponibili, di conseguenza, ad analisi più dettagliate. E' la caratteristica tipica dei software ETL.

**3. Architettura cloud.**

Offerta delle potenzialità del prodotto anche via web, permettendo così di usufruire di tutte le funzionalità senza dover installare nessun componente. E' un ambito che sta prendendo sempre più piede nelle aziende, in quanto comporta dei vantaggi notevoli tra cui, oltre alla non necessità dell'installazione, anche la possibilità di condividere le informazioni con tutti i membri dell'azienda. L'offerta cloud può essere di più tipi: *Infrastructure as a Service* (IaaS), in cui tutta l'architettura hardware viene resa disponibile come servizio, *Platform as a Service* (PaaS), nella quale è fruibile la piattaforma, e *Software as a Service* (SaaS), dove l'applicazione è il servizio offerto al cliente.

**4. Supporto mobile.**

Possibilità di gestire la piattaforma, visualizzare e modificare cruscotti e report su dispositivi mobili come tablet e smartphone. Vengono presi in considerazione anche i vari sistemi operativi sui quali questa funzionalità è supportata (Android, iOS, Windows Phone, etc.).

## **5. Funzionalità di base e Reporting.**

Con “funzionalità di base” si intende la possibilità di esportare gli elaborati in formati comuni, come in file Excel e in Pdf. Inoltre, viene presa in considerazione la gestione da parte del prodotto dei metadati; la presenza di un repository interno per il salvataggio di tutte le informazioni sui dati è uno strumento molto utile per l’azienda e dal quale può trarre ulteriori analisi ed effettuare delle impact analysis, fondamentale per determinare l’impatto di alcune modifiche. Il Reporting rappresenta una delle funzionalità principali e tra le più utilizzate in ambito aziendale.

## **6. Dashboarding, analisi e Data Mining.**

Con questo parametro si vuole indicare quella parte di front end in cui si è più investito negli ultimi anni. Sono comprese la Data visualization, la Data discovery e l’analisi predittiva. A differenza dei report, permettono di analizzare dinamicamente le informazioni, fino ad effettuare simulazioni e previsioni degli andamenti futuri. L’importanza di queste funzionalità è il valore aggiunto che portano in azienda grazie alla scoperta di nuove informazioni.

## **7. Usabilità, Visualization e Self Service.**

Per le PMI è spesso uno degli aspetti più importanti e determinanti nella scelta. Con questa voce si intende l’usabilità del prodotto, quindi la fluidità e i tempi di risposta, l’impaginazione grafica (il numero, la disposizione e la grandezza delle finestre di lavoro che vengono visualizzate) e, soprattutto, la misura in cui l’utente finale sarà in grado di essere autonomo nell’utilizzo. E’ fondamentale, quindi, la semplicità e l’intuitività del prodotto e della sua interfaccia. Questo parametro ha un’importanza notevole, in quanto le PMI non hanno le stesse disponibilità delle grandi imprese da investire nella formazione dei propri dipendenti; di conseguenza, anche la preparazione e le conoscenze tecniche potrebbero essere limitate.

## **8. Analisi dei Big Data.**

E’ un aspetto che ha un’importanza leggermente inferiore rispetto ad altri, in quanto, quello dei Big Data, è un fenomeno che sta prendendo piede negli ultimi anni e che, per adesso, è di interesse principalmente per le grandi aziende. Nonostante ciò, molti prodotti hanno cominciato ad

inserirne il supporto, offrendo anche alle PMI la possibilità di sfruttare questa nuova potenzialità.

## 9. Documentazione e supporto.

Un prodotto con una documentazione approfondita e magari con una community molto attiva, è sicuramente preferibile ad un prodotto con scarse informazioni al suo riguardo. Oltre alla documentazione, molto importante è anche il supporto offerto da tecnici specializzati in caso di necessità.

## 10. Integrazione.

E' uno dei parametri più importanti, a causa del concetto già espresso altre volte nel corso di questo lavoro: la visualizzazione e l'analisi non potranno mai essere efficienti e funzionali se non si basano su una base dati coerente. Per ottenerla è necessario effettuare delle operazioni di ETL (*Extract, Transform, Load*); solo così si può pensare di avere una BI davvero potente.

PARAMETRI SOFTWARE SELECTION		SPIEGAZIONE
1	Infrastruttura e Architettura	> Completezza della suite > Supporto Sistemi Operativi
2	Connettività	> Connettività a sorgenti differenti
3	Architettura Cloud	> Gestione della piattaforma nel cloud > Tipologie di servizi offerti (Saas, PaaS)
4	Supporto Mobile	> Gestione della piattaforma attraverso dispositivi mobile
5	Funzionalità di base e Reporting	> Metadati, Reporting, possibilità di esportare in PDF o Excel...
6	Dashboarding, Analisi e Data Mining	> Data Visualization e Supporto Analisi predittiva e Data Mining
7	Usabilità, Visualization e Self Service	> Usabilità del prodotto > Facilità di utilizzo e di apprendimento > Misura in cui l'utente finale è in grado di essere autonomo
8	Analisi Big Data	> Supporto connettività e analisi Big Data
9	Documentazione e supporto	> Documentazione tecnica e supporto professionale di specialisti
10	Integrazione	> Funzionalità ETL

Tabella 3.2: Parametri software selection

### 3.3.2. Punteggi e valutazioni

Una volta definiti tutti i parametri, per poter confrontare tutti i prodotti, è necessario attribuire sia un'importanza ad ogni singola voce, sia un punteggio in base alle caratteristiche dei vari software. L'importanza che viene attribuita ad ogni parametro è generalmente soggettiva; dipende cioè dalle scelte e dalle filosofie aziendali. Per poter effettuare un'analisi, però, sono stati scelti dei pesi in base a quelle che, generalmente, sono le esigenze delle PMI. La valutazione, invece, dipende dal comportamento e dalle funzionalità del prodotto, oltre che dall'opinione di clienti che già hanno utilizzato il software in esame. Sia ai pesi che alle valutazioni, sono assegnati dei valori che vanno da 0 (zero) a 5 (cinque). Per quanto riguarda i pesi, il valore "0" corrisponde ad un'importanza nulla per la tipologia di aziende presa in considerazione, mentre il valore "5" corrisponde ad un'importanza massima. Per le valutazioni il significato è simile: un punteggio pari a "0" significa che quella determinata *feature* (caratteristica) non è presente per nulla nel prodotto, mentre una valutazione di "5" sta ad indicare che quella determinata funzionalità è presente ed è completa.

<b>PESO</b>	<b>DESCRIZIONE</b>
0	Nessuna importanza per il cliente potenziale
1	
2	
3	
4	
5	Massima importanza per il cliente potenziale

Tabella 3.3: Descrizione pesi

<b>VALUTAZIONE</b>	<b>DESCRIZIONE</b>
0	Non è presente la feature
1	
2	
3	
4	
5	E' presente la feature in modo completo

Tabella 3.4: Descrizione valutazioni

Considerando quindi i parametri selezionati in precedenza, è stato loro attribuito un peso in base all'importanza delle varie caratteristiche dei prodotti, tenendo conto delle esigenze che generalmente le PMI hanno. Ne risulta la seguente tabella:

PARAMETRI SOFTWARE SELECTION		PESO
1	Infrastruttura e Architettura	4
2	Connettività	4
3	Architettura Cloud	2
4	Supporto Mobile	1
5	Funzionalità di base e Reporting	5
6	Dashboarding, Analisi e Data Mining	3
7	Usabilità, Visualization e Self Service	5
8	Analisi Big Data	2
9	Documentazione e supporto	3
10	Integrazione	5

Tabella 3.5: Parametri e pesi utilizzati

Si può osservare come gli aspetti considerati più importanti siano le funzionalità principali in ambito della BI (ETL, Reporting e Dashboarding), insieme alla semplicità del prodotto. Un peso inferiore hanno, invece, il supporto mobile, il cloud e i Big Data che, nonostante sia indubbia la loro utilità e importanza in abito aziendale, sono comunque degli aspetti non fondamentali per una PMI. Ciò non toglie che, se presenti, sono sicuramente un valore aggiunto.

### 3.4. Confronto

Esattamente come per gli RDBMS, anche per i software di BI esistono degli studi che possono aiutare l'IT manager nella scelta della giusta tecnologia. In particolar modo, Gartner offre uno studio molto approfondito, nel quale i prodotti vengono valutati e confrontati in base ai due parametri, descritti anche nel primo paragrafo di questo capitolo: "completezza di visione" e "capacità di esecuzione". Ogni anno, viene pubblicato un documento per ogni tipologia di software; nel caso della BI il documento prende il nome di "*Magic Quadrant for Business Intelligence and Analytics Platform*". Tutte le considerazioni fatte in precedenza sono valide anche in questo caso; è infatti necessario confrontare i prodotti anche in base ai criteri stabiliti dall'azienda e non solo

in base a questo elaborato, che offre comunque un ottimo punto di partenza. Il Magic Quadrant pubblicato il 23 febbraio 2015 è riportato nella figura riportata in seguito.



Figura 3.2: Magic Quadrant di Gartner. (Fonte: Gartner, Febbraio 2015)

Altro studio che può sicuramente aiutare le aziende è quello effettuato da Aberdeen Group, intitolato “*Business Intelligence (BI): Performance Management AXIS*”. Aberdeen AXIS offre una valutazione sui prodotti di business intelligence in base a due criteri, che corrispondono rispettivamente all’asse delle ascisse e a quello delle ordinate: “*Market Readiness*” (letteralmente, prontezza di mercato) e “*Value Delivered*” (letteralmente, valore consegnato). L’analisi viene effettuata in base alle esperienze e alle testimonianze di utenti dei vari prodotti. In base a queste interviste, i vendor vengono classificati in una delle quattro categorie individuate da Aberdeen:

- *Champion*. Vendor che hanno dimostrato una preparazione superiore, sia in termini di potenzialità dei prodotti, sia in termini di assistenza al cliente.
- *Contender*. Vendor che offrono agli utenti una buona soluzione in relazione all'investimento fatto e, allo stesso tempo, la capacità di garantire il supporto tecnico.
- *Competitor*. Sono quelle realtà che si stanno ancora sviluppando e che stanno cercando di trovare il giusto equilibrio tra le potenzialità offerte, il costo, il supporto e la continuità.
- *Challenger*. Vendor che hanno raggiunto una visibilità nel mercato e che hanno la possibilità di soddisfare le esigenze di alcuni utenti in base all'investimento che sono disposti a fare. La prontezza di queste realtà è il loro punto più debole.

Quest'analisi è solo uno dei tanti studi effettuati da Aberdeen Group che, di conseguenza, può davvero essere uno strumento molto importante per un'azienda.

Una volta presi in considerazione degli studi ufficiali, ma inevitabilmente generali, l'azienda deve eseguire un proprio confronto delle potenzialità dei prodotti presi in esame. Di seguito verrà effettuata una valutazione e una comparazione dei prodotti descritti e affrontati nel secondo capitolo, con lo scopo di individuare gli strumenti più adeguati alle esigenze identificate come parametri in precedenza. Il confronto sarà effettuato tra prodotti appartenenti ad una stessa tipologia, anche se qualche considerazione trasversale verrà fatta. La scelta di questa modalità di comparazione dipende dal fatto che due prodotti appartenenti a due tipologie diverse hanno inevitabilmente delle caratteristiche non comuni. Per esempio, un prodotto open source è evidente che non potrà avere tutte le potenzialità che caratterizzano un prodotto storico, così come, viceversa, un prodotto storico non avrà lo stesso costo di uno open source. L'analisi, di conseguenza, non potrebbe essere fatta in maniera equa, in quanto ci sono delle caratteristiche differenti che dipendono dalla natura del prodotto e non dall'investimento e dal lavoro svolto dal vendor.

Ad ogni prodotto è stata assegnata una valutazione per ogni criterio in base alle caratteristiche dichiarate dal sito ufficiale del vendor, alle testimonianze di utenti che hanno adottato il software analizzato e a quelle dimostrate durante l'utilizzo. Per ogni criterio, la valutazione assegnata è stata moltiplicata per il peso, ottenendo così un punteggio.

$$\text{Punteggio parametro} = \text{Valutazione} * \text{Peso}$$

I punteggi di tutti i parametri sono stati successivamente sommati al fine di calcolare un punteggio totale, che servirà per definire se un prodotto potrebbe essere ipoteticamente migliore di altri.

$$\text{Punteggio Totale} = \sum \text{Punteggio parametro}$$

Per ogni criterio è stato calcolato, inoltre, un punteggio relativo, che indica in che percentuale è soddisfatta la determinata *feature*; se un parametro ha un punteggio relativo di 1 (100%) significa che il prodotto ha preso una valutazione di 5 punti su 5, mentre, se ha un punteggio relativo di 0, significa che la valutazione è stata di 0 punti su 5.

$$\begin{aligned} \text{Punteggio Relativo} &= \text{Punteggio} / \text{Max(Valutazione)} \\ &= \text{Punteggio} / 5 \end{aligned}$$

Infine viene calcolata la media pesata

$$\text{Media Pesata} = \text{Punteggio Totale} / \sum \text{Pesi}$$

### 3.4.1. Prodotti Storici

In seguito verranno presentate le tabelle di valutazione dei prodotti storici analizzati assieme ad un grafico a radar rappresentante il punteggio relativo attribuito ai vari parametri.

#### Oracle Corporation

STORICI			ORACLE		
	PARAMETRI	PESO	Valutazione	Punteggio	Punt. Relativo
1	Infrastruttura e Architettura	4	5	20	1
2	Connettività	4	5	20	1
3	Architettura Cloud	2	3	6	0,6
4	Supporto Mobile	1	2	2	0,4
5	Funzionalità di base e Reporting	5	3	15	0,6
6	Dashboarding, Analisi e Data Mining	3	3	9	0,6
7	Usabilità, Visualization e Self Service	5	2	10	0,4
8	Analisi Big Data	2	3	6	0,6
9	Documentazione e supporto	3	5	15	1
10	Integrazione	5	5	25	1
	<b>PUNTEGGIO TOT</b>	<b>34</b>		<b>128</b>	
	<b>MEDIA PESATA</b>			<b>3,8</b>	

Tabella 3.6: Valutazione Oracle

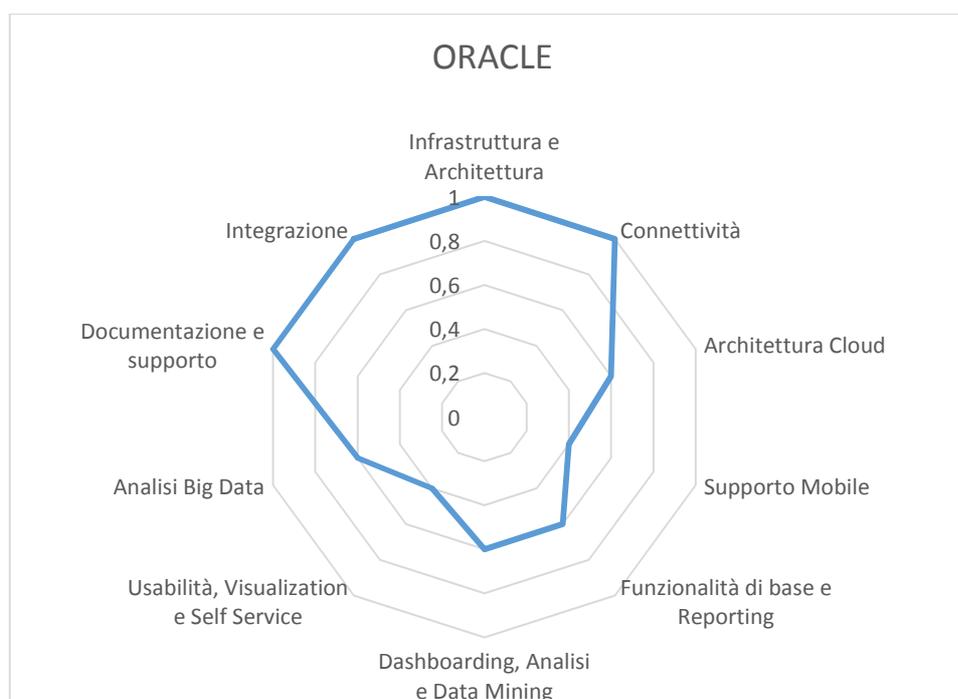


Figura 3.3: Grafico valutazione Oracle

## Microsoft Corporation

STORICI			MICROSOFT		
	PARAMETRI	PESO	Valutazione	Punteggio	Punt. Relativo
1	Infrastruttura e Architettura	4	4	16	0,8
2	Connettività	4	4	16	0,8
3	Architettura Cloud	2	3	6	0,6
4	Supporto Mobile	1	4	4	0,8
5	Funzionalità di base e Reporting	5	4	20	0,8
6	Dashboarding, Analisi e Data Mining	3	3	9	0,6
7	Usabilità, Visualization e Self Service	5	5	25	1
8	Analisi Big Data	2	2	4	0,4
9	Documentazione e supporto	3	4	12	0,8
10	Integrazione	5	4	20	0,8
	<b>PUNTEGGIO TOT</b>	<b>34</b>		<b>132</b>	
	<b>MEDIA PESATA</b>			<b>3,9</b>	

Tabella 3.7: Valutazione Microsoft

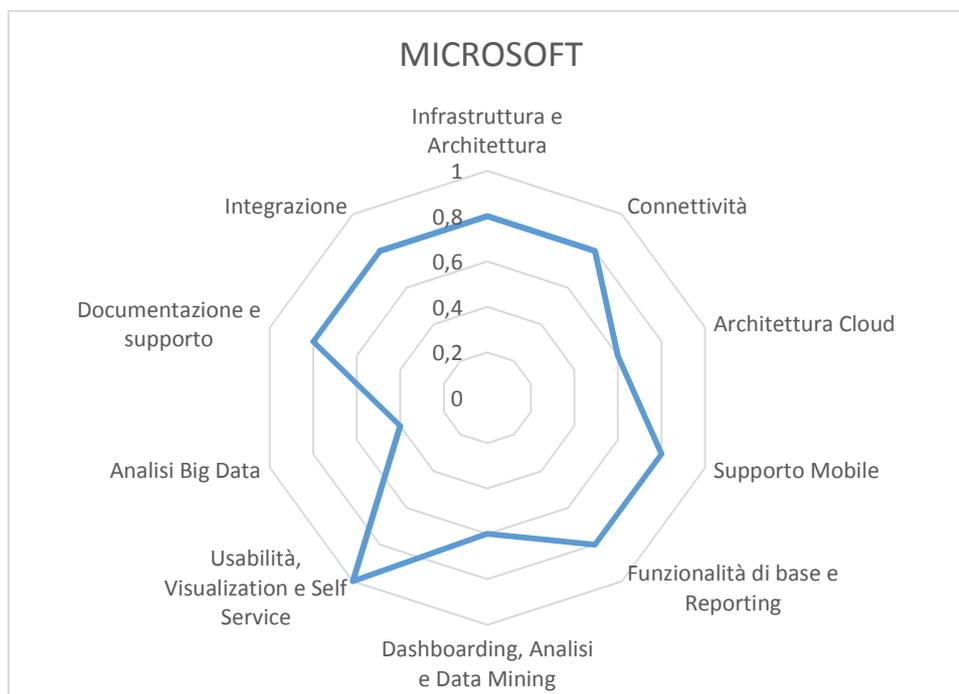


Figura 3.4: Grafico valutazione Microsoft

## IBM

STORICI			IBM		
	PARAMETRI	PESO	Valutazione	Punteggio	Punt. Relativo
1	Infrastruttura e Architettura	4	4	16	0,8
2	Connettività	4	5	20	1
3	Architettura Cloud	2	3	6	0,6
4	Supporto Mobile	1	3	3	0,6
5	Funzionalità di base e Reporting	5	5	25	1
6	Dashboarding, Analisi e Data Mining	3	4	12	0,8
7	Usabilità, Visualization e Self Service	5	3	15	0,6
8	Analisi Big Data	2	3	6	0,6
9	Documentazione e supporto	3	1	3	0,2
10	Integrazione	5	3	15	0,6
<b>PUNTEGGIO TOT</b>		<b>34</b>		<b>121</b>	
<b>MEDIA PESATA</b>				<b>3,6</b>	

Tabella 3.8: Valutazione IBM

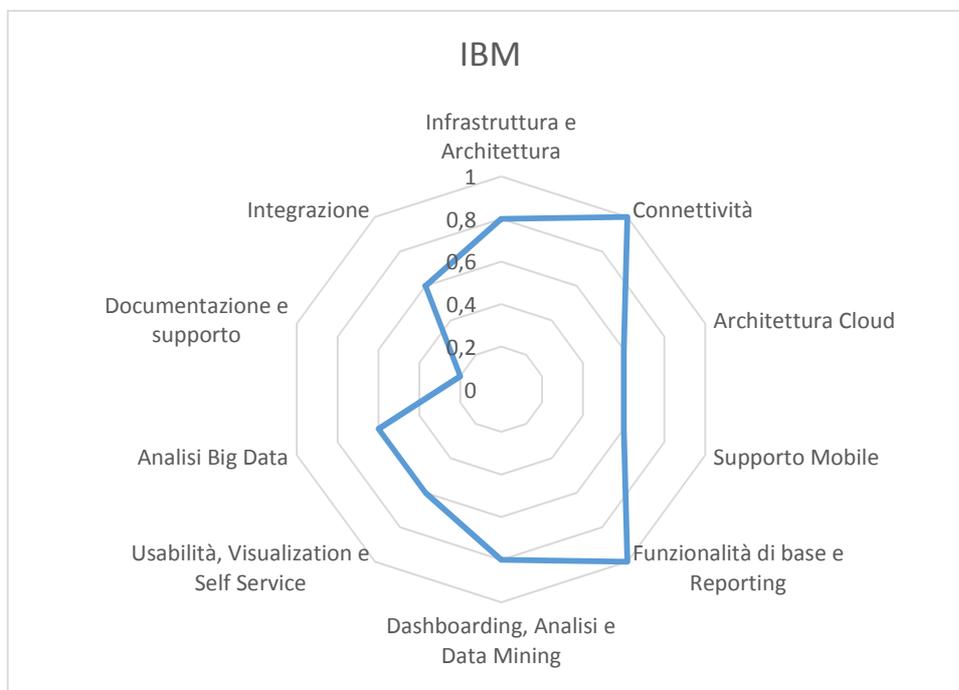


Figura 3.5: Grafico valutazione IBM

## SAP

STORICI			SAP		
	PARAMETRI	PESO	Valutazione	Punteggio	Punt. Relativo
1	Infrastruttura e Architettura	4	4	16	0,8
2	Connettività	4	5	20	1
3	Architettura Cloud	2	3	6	0,6
4	Supporto Mobile	1	3	3	0,6
5	Funzionalità di base e Reporting	5	4	20	0,8
6	Dashboarding, Analisi e Data Mining	3	2	6	0,4
7	Usabilità, Visualization e Self Service	5	2	10	0,4
8	Analisi Big Data	2	2	4	0,4
9	Documentazione e supporto	3	4	12	0,8
10	Integrazione	5	4	20	0,8
	<b>PUNTEGGIO TOT</b>	34		117	
	<b>MEDIA PESATA</b>			3,4	

Tabella 3.9: Valutazione SAP

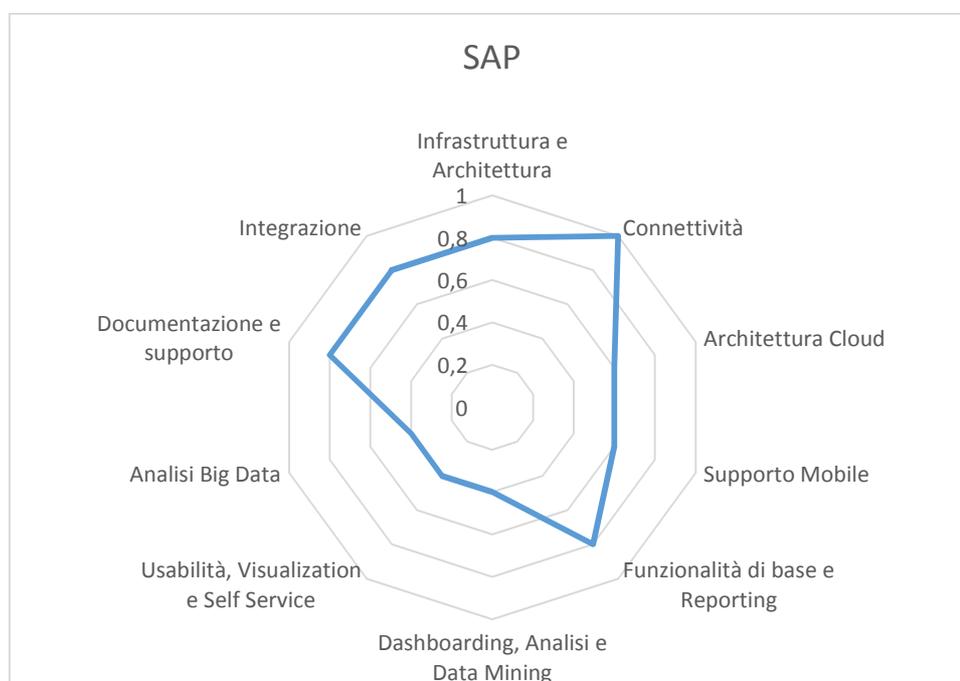


Figura 3.6: Grafico valutazione SAP

## MicroStrategy

STORICI			MICROSTRATEGY		
	PARAMETRI	PESO	Valutazione	Punteggio	Punt. Relativo
1	Infrastruttura e Architettura	4	5	20	1
2	Connettività	4	4	16	0,8
3	Architettura Cloud	2	4	8	0,8
4	Supporto Mobile	1	5	5	1
5	Funzionalità di base e Reporting	5	4	20	0,8
6	Dashboarding, Analisi e Data Mining	3	4	12	0,8
7	Usabilità, Visualization e Self Service	5	3	15	0,6
8	Analisi Big Data	2	2	4	0,4
9	Documentazione e supporto	3	4	12	0,8
10	Integrazione	5	2	10	0,4
<b>PUNTEGGIO TOT</b>		<b>34</b>		<b>122</b>	
<b>MEDIA PESATA</b>				<b>3,6</b>	

Tabella 3.10: Valutazione Microstrategy

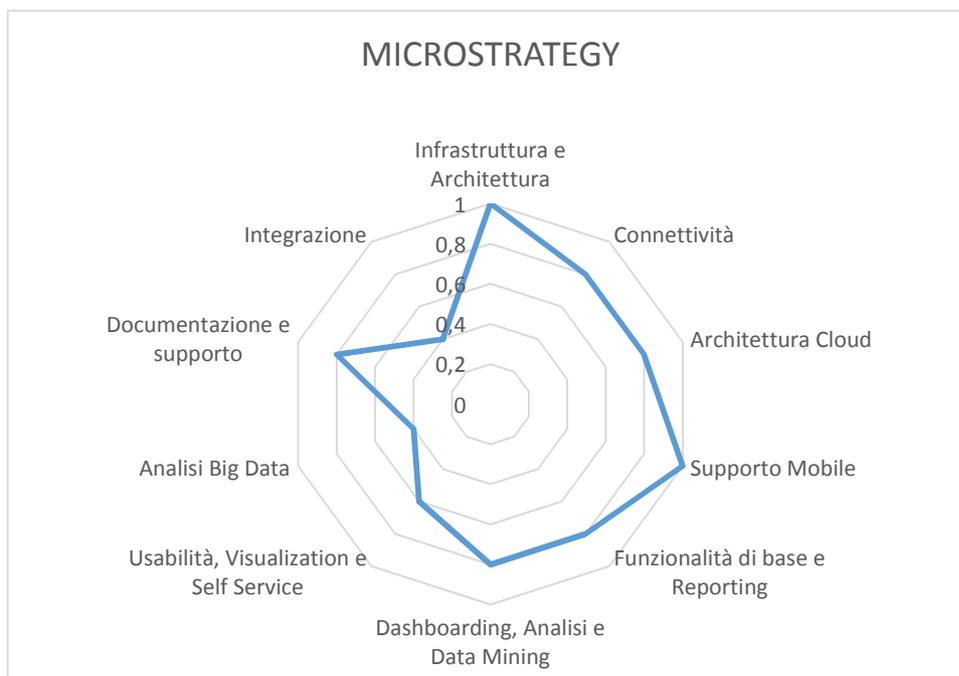


Figura 3.7: Grafico valutazione Microstrategy

Analizzando i punteggi, si possono fare delle considerazioni in merito a questi prodotti storici. Per quanto riguarda la parte di back end, cioè di infrastruttura fisica, Oracle, Microsoft e MicroStrategy offrono delle soluzioni molto solide e consolidate, orientate al DW e ben integrate con gli altri strumenti offerti. MicroStrategy ha ottenuto un punteggio non molto alto a causa dell'assenza nell'offerta di questo vendor di uno strumento di ETL. Fatta eccezione per questo aspetto, risulta essere probabilmente il prodotto complessivamente migliore dopo i due colossi Oracle e Microsoft. Se si analizza invece la parte di front end, IBM, grazie a Cognos, ricopre un ruolo molto importante e offre ai clienti una tra le migliori soluzioni in ambito di reporting, data visualization e anche data discovery. SAP è un prodotto orientato più al mondo gestionale che non alla BI, infatti risulta essere leggermente inferiore rispetto ai concorrenti; elemento positivo di SAP è sicuramente Business Objects dal punto di vista del reporting.

Generalmente, questi prodotti sono caratterizzati dalle maggiori potenzialità offerte nel mercato, accompagnate però anche dai costi più alti. Per questo motivo, sono molto interessanti le proposte offerte da Microsoft, che sono in continuo sviluppo, e ormai, allo stesso livello degli altri prodotti, ma con prezzi leggermente più accessibili e, soprattutto, con una facilità di utilizzo che non si trova in nessun altro dei prodotti storici analizzati. Questo fattore è spesso considerato fondamentale da molti utenti e, senza dubbio, in ambito delle PMI, Microsoft rappresenta un'alternativa molto valida.

### 3.4.2. Prodotti Indipendenti

Stesso procedimento e stessa analisi sono stati adottati anche per i prodotti indipendenti.

#### Qlik

INDIPENDENTI			QLIK		
	PARAMETRI	PESO	Valutazione	Punteggio	Punt. Relativo
1	Infrastruttura e Architettura	4	3	12	0,6
2	Connettività	4	3	12	0,6
3	Architettura Cloud	2	2	4	0,4
4	Supporto Mobile	1	4	4	0,8
5	Funzionalità di base e Reporting	5	2	10	0,4
6	Dashboarding, Analisi e Data Mining	3	5	15	1
7	Usabilità, Visualization e Self Service	5	4	20	0,8
8	Analisi Big Data	2	2	4	0,4
9	Documentazione e supporto	3	3	9	0,6
10	Integrazione	5	2	10	0,4
<b>PUNTEGGIO TOT</b>		<b>34</b>		<b>100</b>	
<b>MEDIA PESATA</b>				<b>2,9</b>	

Tabella 3.11: Valutazione Qlik

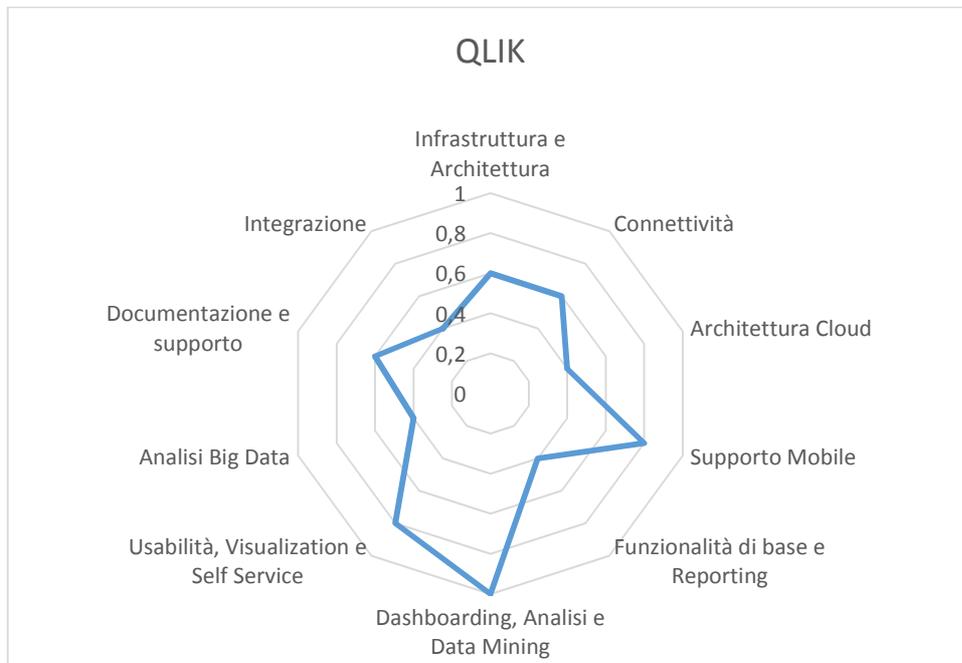


Figura 3.8: Grafico valutazione Qlik

## TIBCO Spotfire

INDIPENDENTI			SPOTFIRE		
	PARAMETRI	PESO	Valutazione	Punteggio	Punt. Relativo
1	Infrastruttura e Architettura	4	2	8	0,4
2	Connettività	4	3	12	0,6
3	Architettura Cloud	2	3	6	0,6
4	Supporto Mobile	1	3	3	0,6
5	Funzionalità di base e Reporting	5	2	10	0,4
6	Dashboarding, Analisi e Data Mining	3	5	15	1
7	Usabilità, Visualization e Self Service	5	5	25	1
8	Analisi Big Data	2	2	4	0,4
9	Documentazione e supporto	3	2	6	0,4
10	Integrazione	5	1	5	0,2
	<b>PUNTEGGIO TOT</b>	<b>34</b>		<b>94</b>	
	<b>MEDIA PESATA</b>			<b>2,8</b>	

Tabella 3.12: Valutazione Spotfire

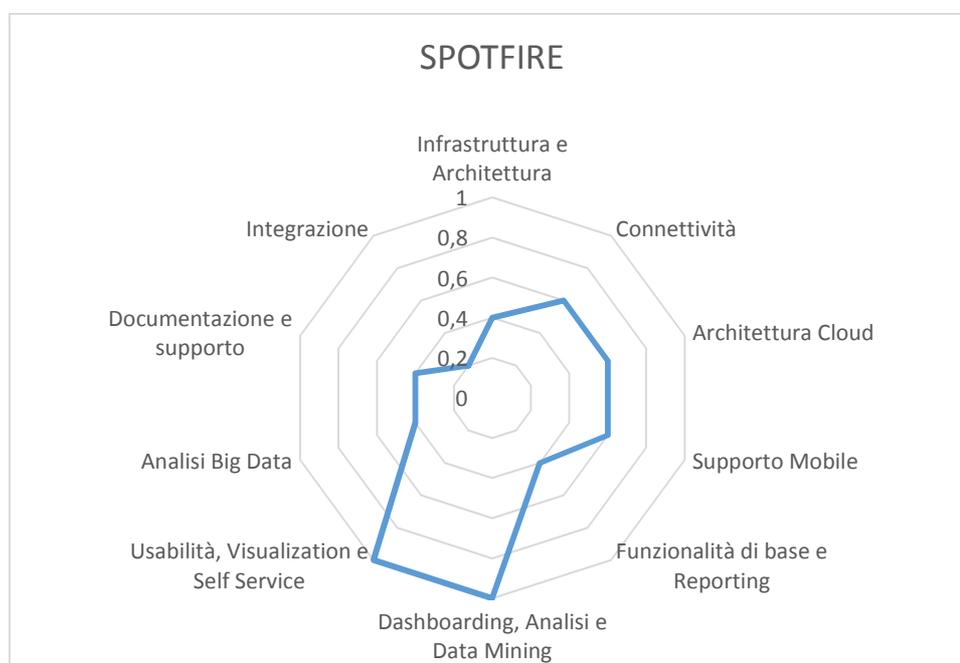


Figura 3.9: Grafico valutazione Spotfire

## Tableau

INDIPENDENTI			TABLEAU		
	PARAMETRI	PESO	Valutazione	Punteggio	Punt. Relativo
1	Infrastruttura e Architettura	4	3	12	0,6
2	Connettività	4	3	12	0,6
3	Architettura Cloud	2	3	6	0,6
4	Supporto Mobile	1	4	4	0,8
5	Funzionalità di base e Reporting	5	1	5	0,2
6	Dashboarding, Analisi e Data Mining	3	5	15	1
7	Usabilità, Visualization e Self Service	5	5	25	1
8	Analisi Big Data	2	2	4	0,4
9	Documentazione e supporto	3	3	9	0,6
10	Integrazione	5	2	10	0,4
	<b>PUNTEGGIO TOT</b>	<b>34</b>		<b>102</b>	
	<b>MEDIA PESATA</b>			<b>3,0</b>	

Tabella 3.13: Valutazione Tableau

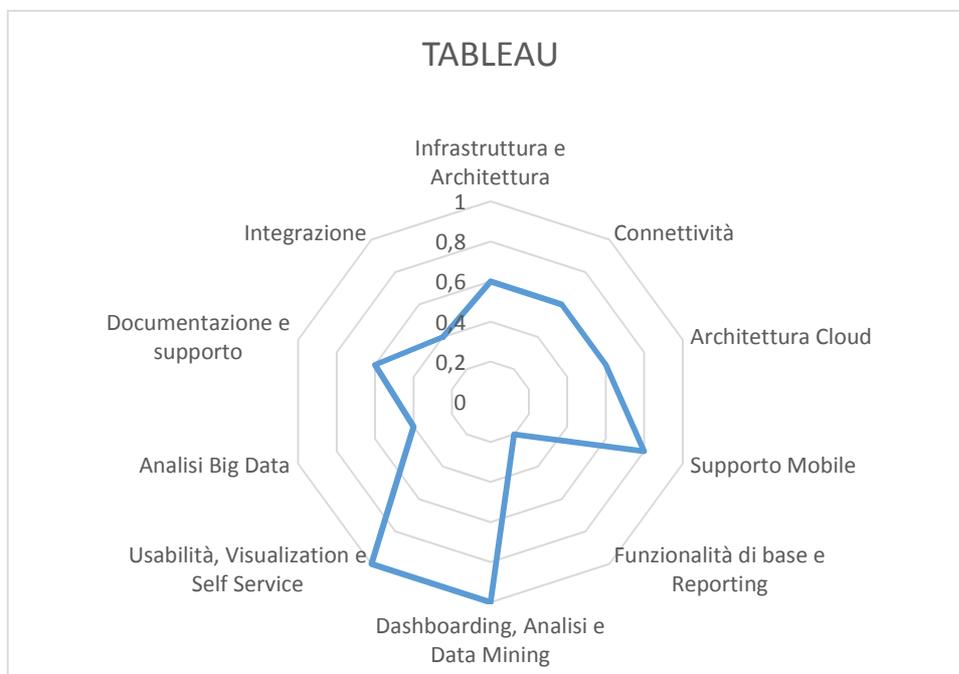


Figura 3.10: Grafico valutazione Tableau

I prodotti indipendenti sono caratterizzati da uno sviluppo delle funzionalità di front end molto importante e, viceversa, da un'attenzione nettamente inferiore a tutta la parte di back end. Di conseguenza, l'analisi si soffermerà sulle caratteristiche che contraddistinguono in maniera positiva questi prodotti. Se si guarda alla completezza della soluzione offerta, l'attenzione si rivolge a Qlik, in quanto, grazie alla nuova piattaforma Qlik Sense, offre una serie di funzionalità che vanno a coprire quasi tutta la BI (le due aree meno sviluppate sono il reporting e l'analisi predittiva). Questo strumento è, inoltre, quello che presenta la maggior velocità, grazie al suo motore interno basato sulla teoria associativa. D'altra parte, però, Tableau è il prodotto che offre le maggiori potenzialità in termini di visualizzazione dei dati ed è più semplice da utilizzare rispetto a Qlik. Più semplice ancora di Tableau, però, è Spotfire, che fa di questo aspetto il suo punto di forza, assieme al supporto di sofisticate analisi matematiche che gli altri due prodotti non prevedono, o prevedono in maniera limitata. Sono tutte e tre delle ottime soluzioni per il front end.

Se l'esigenza dell'azienda, perciò, è la visualizzazione dei dati, essa sarà maggiormente interessata a Tableau; se invece vede nella semplicità di utilizzo la priorità, allora sarà orientata verso Spotfire; se, infine, necessita di uno strumento che offra la possibilità di creare complessi elaborati in maniera grafica per poter visualizzare dati aggregati, allora sceglierà presumibilmente Qlik.

### 3.4.3. Prodotti Open Source

Infine sono stati affrontati i prodotti definiti open source che presentano delle caratteristiche interessanti, considerando che sono generalmente gratuite. Di seguito le tabelle e i grafici di valutazione.

#### **Pentaho**

OPEN SOURCE			PENTAHO		
	PARAMETRI	PESO	Valutazione	Punteggio	Punt. Relativo
1	Infrastruttura e Architettura	4	3	12	0,6
2	Connettività	4	4	16	0,8
3	Architettura Cloud	2	3	6	0,6
4	Supporto Mobile	1	1	1	0,2
5	Funzionalità di base e Reporting	5	3	15	0,6
6	Dashboarding, Analisi e Data Mining	3	2	6	0,4
7	Usabilità, Visualization e Self Service	5	5	25	1
8	Analisi Big Data	2	3	6	0,6
9	Documentazione e supporto	3	1	3	0,2
10	Integrazione	5	4	20	0,8
<b>PUNTEGGIO TOT</b>		<b>34</b>		<b>110</b>	
<b>MEDIA PESATA</b>				<b>3,2</b>	

Tabella 3.14: Valutazione Pentaho

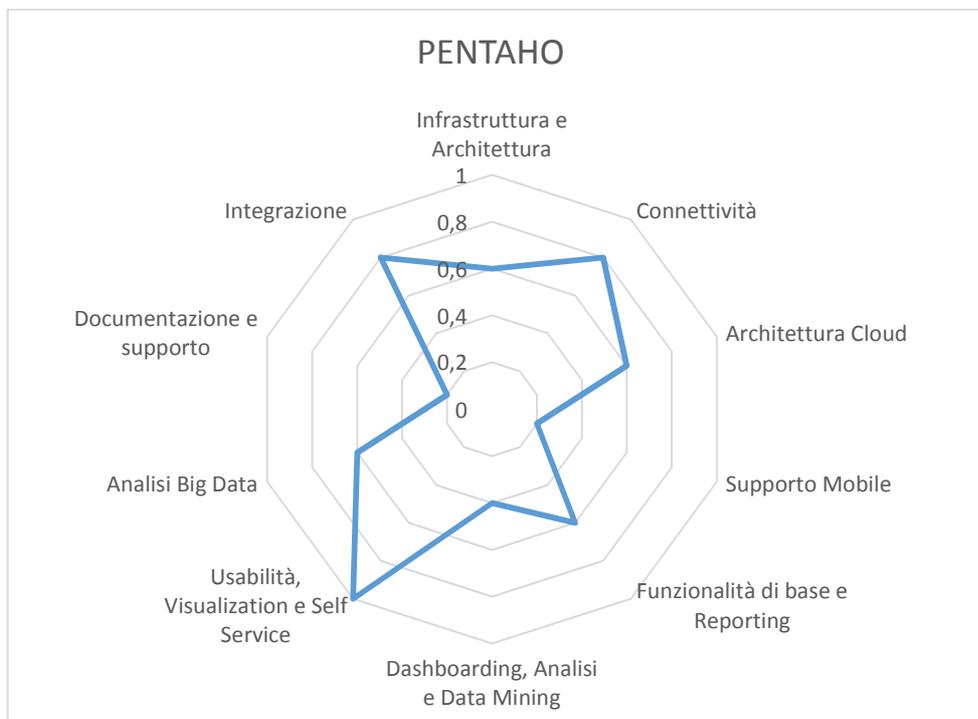


Figura 3.11: Grafico valutazione Pentaho

## TIBCO Jaspersoft BI

OPEN SOURCE			JASPERSOFT		
	PARAMETRI	PESO	Valutazione	Punteggio	Punt. Relativo
1	Infrastruttura e Architettura	4	3	12	0,6
2	Connettività	4	3	12	0,6
3	Architettura Cloud	2	3	6	0,6
4	Supporto Mobile	1	3	3	0,6
5	Funzionalità di base e Reporting	5	5	25	1
6	Dashboarding, Analisi e Data Mining	3	1	3	0,2
7	Usabilità, Visualization e Self Service	5	2	10	0,4
8	Analisi Big Data	2	3	6	0,6
9	Documentazione e supporto	3	3	9	0,6
10	Integrazione	5	3	15	0,6
	<b>PUNTEGGIO TOT</b>	34		101	
	<b>MEDIA PESATA</b>			3,0	

Tabella 3.15: Valutazione Jaspersoft

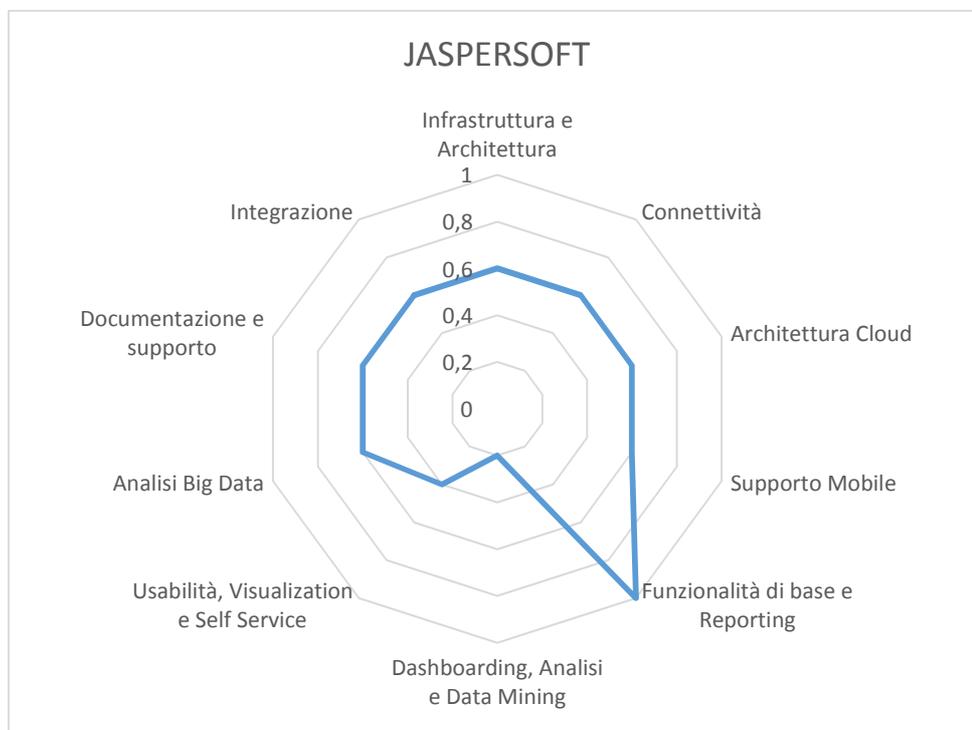


Figura 3.12: Grafico valutazione Jaspersoft

## Eclipse BIRT Project

OPEN SOURCE			BIRT		
	PARAMETRI	PESO	Valutazione	Punteggio	Punt. Relativo
1	Infrastruttura e Architettura	4	2	8	0,4
2	Connettività	4	3	12	0,6
3	Architettura Cloud	2	2	4	0,4
4	Supporto Mobile	1	1	1	0,2
5	Funzionalità di base e Reporting	5	4	20	0,8
6	Dashboarding, Analisi e Data Mining	3	2	6	0,4
7	Usabilità, Visualization e Self Service	5	5	25	1
8	Analisi Big Data	2	3	6	0,6
9	Documentazione e supporto	3	2	6	0,4
10	Integrazione	5	1	5	0,2
	<b>PUNTEGGIO TOT</b>	<b>34</b>		<b>93</b>	
	<b>MEDIA PESATA</b>			<b>2,7</b>	

Tabella 3.16: Valutazione Eclipse BIRT Project

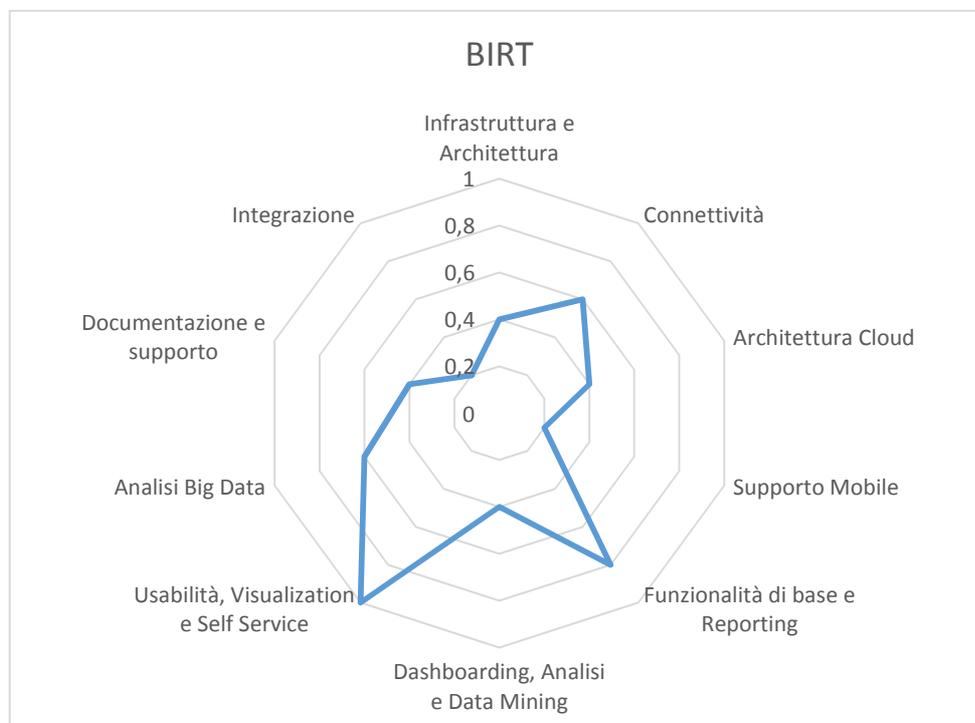


Figura 3.13: Grafico valutazione Eclipse BIRT Project

Tutti e tre questi prodotti sono caratterizzati da un supporto al reporting molto valido e all'altezza, se non superiore, a quelli offerti dagli altri prodotti analizzati in precedenza. In questo ambito, il prodotto migliore risulta essere Jaspersoft, anche se, sia da quanto emerge dalle valutazioni generali fatte da questa analisi, che da altre comparazioni, Pentaho è considerato generalmente migliore rispetto agli altri due prodotti open source. Aspetto che fa la differenza è la parte di back end; Pentaho offre uno strumento ETL (Pentaho Data Integrator) nettamente superiore alle altre soluzioni offerte sia da BIRT che da Jaspersoft e offre agli utenti un'alternativa che si avvicina molto a quelle dei prodotti storici. BIRT risulta essere un prodotto valido soprattutto nella parte di reporting, anche se, messo a confronto con i diretti concorrenti, è caratterizzato da una valutazione complessivamente inferiore, in quanto non offre una soluzione completa come Pentaho.

### **3.5. Alcune soluzioni e suggerimenti**

Questo lavoro non vuole avere la pretesa di presentare la soluzione ottimale per le PMI in ambito dell'infrastruttura informatica, anche perché, vista la diversità delle tipologie di impresa e la varietà dei fattori che guidano gli IT manager nella scelta, non può esistere una soluzione che vada bene per tutte le PMI indistintamente. Quello che si vuole presentare sono alcune delle possibili soluzioni implementabili in un'azienda di dimensioni medio-piccole.

Innanzitutto, è possibile adottare una soluzione *Best of Breed* (BoB), oppure una soluzione *All in One*. Con il termine Best of Breed si vuole indicare un'applicazione software sviluppata per soddisfare un sottoinsieme dei fabbisogni informativi dell'azienda; si tratta quindi di software specializzati in una particolare area della BI, di cui rappresentano una delle soluzioni migliori offerte dal mercato. Adottare una soluzione BoB significa quindi scegliere, per ogni area di interesse, il prodotto che meglio soddisfa le esigenze aziendali, anche se appartenenti a vendor diversi. Questi prodotti prendono il nome di *software dipartimentali*. La filosofia di un sistema BoB comporta degli indubbi vantaggi per quanto riguarda le prestazioni e la competitività di tutto il sistema ma, d'altro canto, comporta dei costi molto elevati che spesso una PMI non può permettersi. Altra problematica che l'azienda, in questo caso, si troverebbe ad affrontare, è l'integrazione di tutte le applicazioni scelte al fine di costituire un sistema informatico efficiente in tutto il suo insieme. In termini pratici, quindi, l'azienda dispone delle migliori funzionalità presenti nel mercato, che devono però essere collegate in modo coerente e prestante. Una soluzione All in One, viceversa, consiste nella scelta di un prodotto che copre, in modo più o meno completo, tutte le aree della BI, senza dover ricorrere a più software

e, quindi, a più licenze. I vantaggi sono, di conseguenza, un costo nettamente inferiore e un'integrazione tra i prodotti ovviamente migliore. Si può notare, però, che nemmeno i prodotti più diffusi e rinomati offrono una soluzione ottimale a tutti i fabbisogni informativi aziendali. La scelta dipende quasi esclusivamente dai vincoli di budget e dalle esigenze prestazionali dell'azienda: una soluzione BoB sarà più costosa, ma più prestante, mentre un prodotto di classe enterprise risulterà meno efficiente, ma anche meno costoso e meno complesso.

Di seguito verranno presentate tre possibili soluzioni in base alle valutazioni eseguite nel capitolo precedente. Ognuna di esse si basa su un concetto ripetuto altre volte in precedenza, secondo cui è fondamentale implementare un'architettura fisica solida e coerente, su cui poi poter appoggiare i vari software front end. Questo significa che, se un'azienda possiede una base dati omogenea e ben organizzata, può pensare di sfruttare le potenzialità di qualsiasi prodotto di BI, mentre se tutto il sistema si basa su dati incoerenti, si può pensare di installare il miglior software per la visualizzazione e l'analisi dei dati, senza però ottenere dei buoni risultati. Di conseguenza, diventa essenziale la scelta della tecnologia database e del tool di ETL.

La prima soluzione si basa su Oracle Database, il miglior RDBMS assieme a Microsoft SQL Server. Per una PMI, risulta difficile pensare di adottare una soluzione All in One di Oracle, sia per quanto riguarda le installazioni, che la gestione in seguito di tutto il sistema. Oltre ai costi delle licenze, infatti, a causa della complessità dei prodotti, l'azienda si troverebbe a dover investire molto tempo e denaro o nella formazione del proprio personale o nella consulenza di professionisti nella gestione di questa tecnologia. Se, quindi, si sceglie di adottare Oracle come RDBMS, una PMI deve pensare ad una soluzione che prevede prodotti di vendor differenti. Nel momento in cui un'azienda investe per questo database, è consigliabile spendere un'ulteriore somma, piuttosto importante, per la licenza di Oracle Data Integrator. ODI è sicuramente uno dei migliori strumenti di ETL presenti nel mercato e permette di ottenere una base dati pulita e ottimizzata. Visto l'importante investimento ipotizzato per la parte di back end, a questo punto l'azienda potrebbe pensare di adottare dei prodotti open source e/o indipendenti per tutta la parte di front end. Una possibile soluzione per quanto riguarda il reporting potrebbe essere Jaspersoft BI, la piattaforma offerta da TIBCO. E' un prodotto molto diffuso e con funzionalità di reporting pari ad altri prodotti considerati di fascia superiore, ma con un prezzo nullo (nel caso della versione Community) o comunque contenuto (per quanto riguarda invece la versione Enterprise). Per molte PMI, però, potrebbe risultare problematica la gestione di tre tecnologie differenti. Per questo motivo, se l'azienda ha le risorse per farlo, potrebbe acquistare Oracle Business Intelligence che permetterebbe

un'integrazione maggiore con tutta la parte di back end e una copertura ottimale delle funzionalità tipiche del front end. Per coprire infine la parte di data visualization e data discovery, la scelta ricade in uno dei due prodotti leader in questo settore: Qlik Sense e Tableau. Il primo è caratterizzato da una grande ottimizzazione e, quindi, rapidità nei tempi di risposta, mentre il secondo dalle migliori funzionalità per la visualizzazione dei dati. Per quelle che sono le esigenze di una PMI, probabilmente è più adeguato Tableau, strumento più semplice da utilizzare e da gestire e caratterizzato da un costo presumibilmente inferiore (dipende da molti fattori).

La seconda soluzione comprende un sistema All in One basato MicroStrategy. Le aziende che hanno scelto questo prodotto lo hanno fatto per l'architettura che offre; grazie all'architettura a tre livelli e al motore ROLAP interno è uno dei prodotti più rapidi e agili, che fa dell'unicità della fonte dei dati per il front end il suo punto di forza. MicroStrategy, infatti, integra tutti i dati delle varie sorgenti nel suo motore interno sul quale basa poi tutti i tool di front end. Per quanto riguarda il reporting, la creazione di dashboard e anche il data mining, offre MicroStrategy Desktop, supportato sia dal sistema operativo di Apple sia da Windows. Inoltre, con la versione Web e Mobile, permette di gestire e di sfruttare tutte le funzionalità offerte tramite il browser e per mezzo di dispositivi mobili. In realtà, questa soluzione necessita di uno strumento di ETL di un altro vendor, in quanto l'offerta di MicroStrategy non ne prevede uno. In questo caso, un'ottima soluzione potrebbe essere l'utilizzo di Kettle, il tool open source di ETL offerto da Pentaho. Questo prodotto comprende ottime funzionalità e permette di caricare il database interno in modo coerente e, soprattutto, senza investimenti economici per la licenza. Si possono così evitare ulteriori spese, in quanto uno degli aspetti non proprio positivi di MicroStrategy è il costo piuttosto elevato; è infatti una soluzione adottata principalmente dalle grandi imprese. Se, però, lo si utilizza come soluzione All in One integrandolo con Kettle, potrebbe risultare un'ottima ipotesi per le medie imprese.

La terza alternativa è una soluzione spesso sottovalutata dalle aziende: Microsoft. Questa società non ha una reputazione alla pari degli altri prodotti storici in ambito aziendale, ma questo pregiudizio, che probabilmente in passato poteva essere fondato, con gli ultimi investimenti deve essere rivisto. Microsoft, infatti, offre delle opportunità molto importanti a costi inferiori rispetto ai diretti concorrenti nel mercato. Come nella seconda soluzione, anche in questo caso ipotizziamo la creazione di un sistema All in One, questa volta però con Microsoft. Come già descritto nel secondo capitolo, Microsoft basa tutti gli strumenti sul proprio RDBMS, cioè Microsoft SQL Server, diretto concorrente di Oracle. La principale funzionalità offerta da questo database è SQL Server Integration Services, lo strumento di back end di Microsoft, molto semplice da

utilizzare e allo stesso tempo molto potente. Per la parte di reporting è presente SQL Server Reporting Services, strumento che si basa su un'astrazione logica della struttura fisica del DB, creata attraverso Microsoft BI Semantic Model. Il punto debole di MS, storicamente, è sempre stata la parte di data visualization e l'analisi dei dati. Grazie però a Microsoft Power BI, sta riuscendo a colmare quel gap rispetto agli altri prodotti. Power BI è uno strumento che permette la creazione e la gestione, anche via mobile, sia di report che di dashboard e che supporta la visualizzazione sostanzialmente in real time, attraverso una serie di strutture grafiche molto sofisticate. Microsoft è un vero e proprio prodotto di classe enterprise che consente all'azienda di adottare un unico prodotto per tutte le esigenze. Ulteriore aspetto che rende questa soluzione una tra le più importanti realtà per le PMI è la semplicità di utilizzo dei prodotti messi a disposizione, che non richiedono particolari conoscenze tecniche come, per esempio, accade per Oracle. Per tutte queste caratteristiche, Microsoft non può non essere considerata come possibile soluzione, soprattutto in ambito delle PMI.

Quelle presentate sono solo alcune di tutte le possibili soluzioni. I prodotti presentati nel secondo capitolo e non racchiusi nelle alternative scelte sono comunque dei buoni prodotti. I prodotti open source, per esempio, sono delle ottime alternative agli altri prodotti, soprattutto in ambito delle piccole imprese che non possono permettersi degli investimenti importanti. Spesso, però, le aziende preferiscono spendere una cifra superiore per avere un prodotto consolidato e con un supporto al cliente garantito; questi aspetti penalizzano leggermente la diffusione dei prodotti open source a livello aziendale. Altra considerazione da fare riguarda i prodotti storici non citati; essi offrono quasi sempre le maggiori funzionalità, ma spesso queste sono contenute in soluzioni che possono permettersi solamente le grandi imprese, come accade per esempio per IBM. Stesso discorso può essere fatto per Qlik, utilizzato più nelle grandi imprese che nelle PMI.

Un'ultima nota è giusto farla sulla parte di analisi predittiva. Come approfondito nel primo capitolo, la predictive analytics sta sempre più diffondendosi anche all'interno delle PMI, anche se non sono ancora molte le aziende disposte ad investire tempo e risorse in questa tipologia di analisi. Inoltre, non sono tanti i vendor che offrono strumenti adeguati e sufficientemente sviluppati per queste esigenze. Di conseguenza, nelle soluzioni proposte non sono stati esplicitamente considerati alcuni prodotti; questo non toglie che, se un'azienda ha le possibilità e decide di investire in questo ambito, può trovare dei prodotti che le permettano di trarre dei benefici da questi investimenti. Il principale prodotto analizzato che soddisfa queste esigenze è sicuramente IBM SPSS.

<b>SOLUZIONE 1</b>			
<b>ARCHITETTURA</b>	<b>ETL</b>	<b>REPORTING</b>	<b>DASHBOARDING E ANALISI</b>
Oracle Database	Oracle Data Integrator	Oracle Business Intelligence	Tableau

<b>SOLUZIONE 2</b>			
<b>ARCHITETTURA</b>	<b>ETL</b>	<b>REPORTING</b>	<b>DASHBOARDING E ANALISI</b>
MicroStrategy Intelligence Server	Pentaho Data Integrator (Kettle)	MicroStrategy (Desktop, Web e Mobile)	MicroStrategy (Desktop, Web e Mobile)

<b>SOLUZIONE 3</b>			
<b>ARCHITETTURA</b>	<b>ETL</b>	<b>REPORTING</b>	<b>DASHBOARDING E ANALISI</b>
Microsoft SQL Server	SQL Server Integration Services	SQL Server Reporting Services Microsoft Power BI	Microsoft Power BI

Tabella 3.17: Tabella riassuntiva delle tre soluzioni proposte

# CAPITOLO 4

## Conclusioni

Lo scopo di questo lavoro è quello di ripercorrere il processo che un'azienda deve fare per progettare e implementare un sistema informativo che possa permetterle di utilizzare l'enorme quantità di dati che si trova a gestire per poterli trasformare in informazioni utili al processo decisionale. L'importanza di un sistema informativo ben sviluppato, già discusso nel primo capitolo, deve essere ben nota alle aziende, in quanto permette loro di aumentare le conoscenze e le esperienze che possono fare la differenza all'interno di un mercato molto affollato e agguerrito.

La prima fase di questo processo consiste nella definizione delle esigenze interne e dei vincoli aziendali che dipendono dalla singola realtà e, di conseguenza, di difficile definizione per le PMI in generale.

Le due fasi su cui si è posta maggiormente l'attenzione sono l'analisi del mercato e la comparazione dei prodotti individuati. Nella panoramica del mercato si sono suddivisi i prodotti in tre diverse tipologie in base alla loro diffusione, fatturato e licenze. Sono stati selezionati i prodotti generalmente più diffusi e utilizzati in ambito aziendale, che corrispondono conseguentemente agli strumenti più affidabili e con una prospettiva di sviluppo e di permanenza nel mercato di lungo termine. Sono poche le aziende disposte a rischiare nell'investire su un prodotto nuovo, anche se, magari, presenta delle buone caratteristiche; la solidità del sistema informativo è un aspetto fondamentale per tutte le imprese. Le caratteristiche presentate per ogni prodotto sono state individuate grazie ai siti ufficiali dei vendor, a testimonianze di persone che utilizzano o hanno utilizzato in ambito aziendale i vari software e a test pratici delle funzionalità offerte. E' importante analizzare criticamente le caratteristiche descritte nei siti ufficiali, in quanto tendono ovviamente a presentare i propri prodotti come i più completi e i migliori. Nel momento in cui si ha una visione chiara delle principali soluzioni offerte dal mercato, l'azienda deve individuare dei parametri per la comparazione e attribuire ad ognuno un'importanza. In questo lavoro sono stati individuati dieci parametri a cui è stato attribuito un peso in base alla rilevanza in ambito PMI. Ad ogni criterio, successivamente, è stata assegnata una valutazione in relazione alla completezza delle soluzioni offerte dal prodotto per soddisfare le esigenze individuate dai parametri. Al termine dell'analisi e della

valutazione, ogni prodotto ha ottenuto un punteggio totale, dato dalla somma di tutte le valutazioni moltiplicate per i pesi, e una valutazione complessiva (punteggio che può andare da 0 a 5), data dal rapporto tra il punteggio totale e la somma di tutti i pesi.

I risultati ottenuti da quest'analisi sono per la maggior parte in linea con le valutazioni e i confronti effettuati da altre società, come, ad esempio, il lavoro pubblicato di anno in anno da Gartner. Non mancano, però, degli aspetti che differiscono da quest'analisi e anche dall'idea che molte aziende hanno. Ad esempio, Microsoft ha ottenuto il punteggio più alto rispetto a tutti gli altri prodotti. Al giorno d'oggi, Microsoft è una delle migliori soluzioni per le PMI perché comprende costi importanti, ma non tra i più elevati, e una semplicità che nessun altro prodotto della stessa fascia di mercato riesce ad offrire. Viceversa, sia Tableau che Qlik hanno ottenuto un punteggio che non rispecchia la loro posizione da leader indiscussi del Magic Quadrant di Gartner. Ciò è dovuto al fatto che questo lavoro considera tutte le varie funzionalità che caratterizzano la BI e, di conseguenza, nonostante l'indiscusso primato di questi due prodotti nella data visualization e nella data discovery, hanno ottenuto una valutazione non troppo elevata negli altri ambiti dove sono meno sviluppati.

A questo punto, grazie a questo lavoro e ad altri elaborati come quello di Gartner, l'azienda ha a disposizione degli strumenti che possono, e devono, aiutarla a individuare i prodotti più adatti a ricoprire tutte le esigenze informative nei vari ambiti aziendali. Nel terzo capitolo sono state presentate tre possibili combinazioni di prodotti che potrebbero costituire l'infrastruttura del sistema informativo di una PMI. Ovviamente, non sono le uniche soluzioni possibili, ma sono il risultato dell'intersezione tra l'analisi svolta e i vincoli che, generalmente, ogni realtà presenta e si impone. Questo lavoro non pretende di offrire una soluzione che possa essere ritenuta la migliore in assoluto, anche perché questa non può esistere; esiste soltanto la miglior soluzione attuabile in ogni singola realtà e questo elaborato vuole dare uno strumento in più alle aziende, affinché possano arrivare a perseguirla.

## Bibliografia

- Bonollo G., Guzzonato L., 2013, *Sistemi Informativi. Modelli, strumenti e applicazioni*, AMON edizioni.
- Eclipse BIRT Project, official website (<http://www.eclipse.org/birt>)
- Golfarelli M., Rizzi S., 2006, *Data Warehouse. Teoria e pratica della progettazione*, McGraw-Hill Education; 2 edizione.
- Hatch D., Lock M., 2009, "*Business Intelligence (BI): Performance Management AXIS*", Aberdeen Group, Inc.
- IBM, official website (<http://www.ibm.com>)
- Infodata, 2 Novembre 2015, "*Big Data, tutti i numeri che contano*", Il Sole 24 ORE
- Jaspersoft Business Intelligence Software, official website (<https://www.jaspersoft.com>)
- Microsoft, official website (<https://docs.oracle.com>)
- MicroStrategy official website (<http://www.microstrategy.com>)
- Oracle, Integrate Cloud Application and Platform Services, official website (<http://www.oracle.com>)
- Oracle Documentation (<https://docs.oracle.com>)
- Pasini P., Perego A., 2012, "*Big Data: nuove fonti di conoscenza aziendale e nuovi modelli di management*", SDA Bocconi School of Management (<http://www.sdabocconi.it>).
- Pentaho | Data integration and Business Analytics, official website (<http://www.pentaho.com>)
- Qlik: Business Intelligence | Data visualization tool, official website (<http://www.qlik.com>)
- Sallam R. L., Hostmann B., Schlegel K., Tapadinhas J., Parenteau J., Oestreich T. W., 2015, "*Magic Quadrant for Business Intelligence and Analytics Platforms*", Technology Research, Gartner Inc.

- SAP Software solution | Technology and Applications, official website (<http://go.sap.com>)
- Tableau Software: Business Intelligence and Analytics, official website (<http://www.tableau.com>)
- TIBCO Spotfire – Business Intelligence Analytics Software, official website (<http://spotfire.tibco.com>)