

**POLITECNICO DI MILANO**

**Scuola di Ingegneria Industriale e dell'Informazione**

**Corso di Laurea Magistrale  
in Ingegneria Gestionale**



**ANALISI COMPARATIVA DEGLI APPROCCI AI BIG DATA ANALYTICS NELLE FUNZIONI  
AZIENDALI**

Tesi di Laurea di:

Paola Monti Matr. 818866

Relatore: Carlo Vercellis

Correlatore: Alessandro Piva

*Anno accademico 2015-2016*

# INDICE

<b>INDICE FIGURE</b>	6
<b>INDICE TABELLE</b>	7
<b>INDICE GRAFICI</b>	8
 <b>SOMMARIO</b>	 10
Sommaio	10
Letteratura di riferimento	11
Obiettivi del lavoro	13
Metodologia di ricerca	14
Risultati ottenuti	15
Conclusioni	18
 <b>ABSTRACT</b>	 20
 <b>INTRODUZIONE</b>	 21
 <b>1 L'ERA DEI BIG DATA</b>	 22
1.1 Cosa sono i Big Data?	22
1.2 Benefici conseguibili dall'utilizzo dei Big Data	23
1.3 Barriere all'utilizzo dei Big Data	27
1.4 Tecniche e tecnologie per l'analisi dei Big Data	31
1.5 Tendenze in atto	38
1.6 L'impatto dei Big Data in quattro settori	39
1.6.1 Sanità	39
1.6.2 Pubblica Amministrazione	44
1.6.3 Manufacturing	46
1.6.4 Retail	49
1.7 Governance dei sistemi BDA&BI	53
1.7.1 Data Scientist	53
1.7.2 Chief Data Officer	54
1.8 Azioni e comportamenti da implementare per riuscire a sfruttare i Big Data	57

1.9	Analisi delle survey .....	61
<b>2</b>	<b>L'IMPATTO DEI BIG DATA SUL MARKETING E SULLA CUSTOMER EXPERIENCE .....</b>	<b>64</b>
2.1	Marketing Analytics .....	64
2.1.1	Tipologie di dati raccolti.....	65
2.1.2	Dall'approccio tradizionale all'approccio predittivo .....	65
2.1.3	Benefici conseguibili dall'utilizzo dei Big Data nel Marketing.....	67
2.1.4	Passi da seguire per riuscire a sfruttare i Marketing Analytics .....	68
2.1.5	Ostacoli da superare per sfruttare i Marketing Analytics .....	71
2.2	Social Analytics.....	72
2.2.1	Livelli di maturità delle iniziative di Social Analytics .....	73
2.2.2	Iniziative di Social Analytics .....	74
2.2.3	Sfide da affrontare per sfruttare i Social Analytics .....	78
2.2.4	Tendenze future.....	79
2.2.5	Implicazioni sul Marketing .....	80
2.3	Principali ambiti dei progetti di Big Data nel marketing.....	81
2.3.1	Direct e Digital Marketing .....	81
2.3.2	Customer Micro-segmentation.....	83
2.3.3	Price Optimization.....	84
2.3.4	Location-based Marketing .....	86
2.3.5	In-store Analysis.....	89
2.3.6	Cross-Selling/Up-Selling .....	90
2.3.7	Altri progetti.....	90
2.3.8	Sintesi dei benefici .....	91
2.4	Impatto dei Big Data sulla Customer Experience.....	92
2.5	Survey analysis .....	93
<b>3</b>	<b>STATO DELL'OFFERTA.....</b>	<b>96</b>
3.1	Contesto di riferimento .....	96
3.2	Magic Quadrant for Advanced Analytics .....	97
<b>4</b>	<b>METODOLOGIA DI RICERCA .....</b>	<b>102</b>
4.1	Introduzione.....	102

4.2	Osservatorio Big Data Analytics & Business Intelligence .....	102
4.3	Percorso di ricerca .....	103
4.3.1	Analisi della letteratura.....	103
4.3.2	Obiettivi della ricerca .....	104
4.3.3	Rilevazione e analisi dei dati .....	105
4.3.4	Stesura dei casi di studio.....	109
4.3.5	Workshop.....	111
<b>5</b>	<b>ANALISI DEGLI APPROCCI AI BDA DELLE FUNZIONI AZIENDALI E DEL SUPPORTO FORNITO DALL'OFFERTA .....</b>	<b>112</b>
5.1	Livello di maturità delle aziende ai Big Data Analytics.....	112
5.2	Livello di maturità delle funzioni aziendali ai Big Data Analytics.....	115
5.3	Il ruolo delle startup.....	123
5.3.1	Dati anagrafici .....	124
5.3.2	Dimensioni di analisi .....	126
5.3.3	Le startup italiane .....	131
5.4	Supporto dell'offerta di mercato alle funzioni aziendali nell'ambito dei Big Data Analytics.....	134
5.5	Ambiti progettuali dei bda delle funzioni aziendali mature .....	141
5.5.1	Marketing e Vendite .....	141
5.5.2	Amministrazione, Finanza e Controllo .....	142
5.5.3	Acquisti.....	142
5.5.4	Logistica e Produzione .....	143
5.6	Conclusione.....	143
<b>6</b>	<b>CASI DI STUDIO .....</b>	<b>146</b>
6.1	Advice Group.....	150
6.2	Azienda X.....	152
6.3	Azienda Y.....	154
6.4	Comune di Brescia .....	157
6.5	Eni .....	159
6.6	Jobrapido .....	162
6.7	Lastminute.com .....	164
6.8	Venere.com.....	166

<b>BIBLIOGRAFIA .....</b>	<b>169</b>
---------------------------	------------

## INDICE FIGURE

<i>Figura 1: Le 5V dei Big Data (Google, 2014)</i> .....	23
<i>Figura 2: Dati generati in un minuto (Osservatorio Big Data Analytics &amp; Business Intelligence, maggio 2015)</i> .....	24
<i>Figura 3: Architettura concettuale per l'analisi dei Big Data (Raghupathi W., 2014)</i> .....	36
<i>Figura 4: Programma di digitalizzazione di regole e regolamenti implementato dal Governo Tedesco (Bikar P., 2015)</i> .....	45
<i>Figura 5: Framework Cyber-Physical Systems per le macchine self-aware e self-maintenance(Lee J., 2014)</i> .....	47
<i>Figura 6: Set di competenze che deve avere il CDO per guidare il cambiamento, la diffusione di nuove idee e la collaborazione all'interno di tutta l'organizzazione (Van Rijmenam M., Chief Data Officer Profile, 2015)</i> .....	56
<i>Figura 7: Fattori alla base della necessità dei Marketing Analytics</i> .....	65
<i>Figura 8: Approccio classico e predittivo all'analisi del Marketing mix (Tarka P., 2014)</i> .....	66
<i>Figura 9: Social Media (Google, 2014)</i> .....	72
<i>Figura 10: Impatto della Social Intelligence sul Marketing(Bellini L., 2012)</i> .....	80
<i>Figura 11: Utilizzo dei dispositivi dal 2013 al 2018 (O'Brien B., 2014)</i> .....	86
<i>Figura 12: Il classico customer journey e il viaggio accelerato di fedeltà dei clienti(Edelman D., 2015)</i> .....	92
<i>Figura 13: Magic Quadrant for Advanced Analytics 2015 (Gartner, 2015)</i> .....	98
<i>Figura 14: Passi determinanti il percorso di Ricerca</i> .....	103

# INDICE TABELLE

<i>Tabella 1: Barriere all'utilizzo dei Big Data .....</i>	<i>31</i>
<i>Tabella 2: Tecniche e tecnologie per l'analisi dei Big Data .....</i>	<i>35</i>
<i>Tabella 3: Benefici conseguibili dall'utilizzo dei Big Data nel settore sanitario .....</i>	<i>43</i>
<i>Tabella 4: Benefici derivanti dallo sfruttamento dei Big Data nelle diverse parti della catena del valore manifatturiera.....</i>	<i>49</i>
<i>Tabella 5: Approcci analitici per lo sfruttamento dei Marketing Analytics.....</i>	<i>70</i>
<i>Tabella 6: Social Media Intelligence Framework (BIT BANG, 2014) .....</i>	
<i>Tabella 7: Iniziative di Social Analytics.....</i>	<i>78</i>
<i>Tabella 8: Cruscotto indicatori per la valutazione del livello di maturità delle funzioni aziendali.....</i>	<i>116</i>
<i>Tabella 9: Età e profilo di competenze dei fondatori delle startup italiane.....</i>	<i>133</i>
<i>Tabella 10: Livello dell'offerta di mercato alle funzioni aziendali nell'ambito dei BDA.....</i>	<i>134</i>
<i>Tabella 11: Livello della domanda di BDA da parte delle funzioni aziendali.....</i>	<i>136</i>

## INDICE GRAFICI

<i>Grafico 1: Posizionamento delle funzioni aziendali rispetto nella matrice livello di domanda /livello di offerta.....</i>	<i>16</i>
<i>Grafico 2: Big Data Analytics.....</i>	<i>36</i>
<i>Grafico 3: Benefici conseguibili dall'utilizzo dei Big Data nel Marketing.....</i>	<i>67</i>
<i>Grafico 4: Barriere allo sfruttamento dei Marketing Analytics.....</i>	<i>70</i>
<i>Grafico 5: Livelli di maturità delle iniziative di Social Analytics.....</i>	<i>72</i>
<i>Grafico 6: Sintesi dei benefici apportati dai Big Data in ambito Marketing.....</i>	<i>90</i>
<i>Grafico 7: Scomposizione campione survey c-level per ruolo organizzativo.....</i>	<i>105</i>
<i>Grafico 8: Schema del percorso di ricerca seguito.....</i>	<i>109</i>
<i>Grafico 9: Approccio delle organizzazioni alla gestione dei Big Data.....</i>	<i>112</i>
<i>Grafico 10: Matrice stato di adozione dei Big Data Analytics e approccio seguito dalle organizzazioni.....</i>	<i>113</i>
<i>Grafico 11: Scomposizione delle startup internazionali per anno di nascita.....</i>	<i>123</i>
<i>Grafico 12: Scomposizione delle startup internazionali per industry.....</i>	<i>125</i>
<i>Grafico 13: Suddivisione delle startup internazionali in base finanziamenti ricevuti dalle dal 2012.....</i>	<i>125</i>
<i>Grafico 14: Scomposizione delle startup internazionali per tipologia di dati considerati .....</i>	<i>126</i>
<i>Grafico 15: Scomposizione delle startup internazionali per origine di dati analizzati .....</i>	<i>126</i>
<i>Grafico 16: Scomposizione delle startup internazionali per tipologia e origine di dati analizzati .....</i>	<i>127</i>
<i>Grafico 17: Scomposizione delle startup internazionali per ambiti applicativi.....</i>	<i>130</i>
<i>Grafico 18: Scomposizione delle startup internazionali focalizzate sul settore retail per ambiti applicativi.....</i>	<i>131</i>



<i>Grafico 19: Scomposizione delle startup italiane per zona geografica.....</i>	<i>132</i>
<i>Grafico 20: Scomposizione delle startup italiane per regione.....</i>	<i>132</i>
<i>Grafico 21: Suddivisione dei finanziamenti ricevuti dalle startup italiane a partire dal 2012 per regione.....</i>	<i>133</i>
<i>Grafico 22: Posizionamento delle funzioni aziendali rispetto nella matrice livello di domanda /livello di offerta.....</i>	<i>139</i>
<i>Grafico 23: Presenza attuale e interesse prospettico verso progetti Big Data nel Marketing .....</i>	<i>142</i>

# SOMMARIO

## SOMMARIO

Il presente lavoro di tesina si colloca nell'ambito di ricerca dei Big Data, il nuovo trend che porterà a grandi cambiamenti nel modo di lavorare delle organizzazioni, le quali potranno trarre enormi benefici tra cui un incremento della competitività, della redditività, della tempestività e dell'efficacia dei processi decisionali. L'elaborato si pone l'obiettivo di indagare il livello di maturità delle aziende e in modo approfondito quello delle funzioni aziendali, analizzando i progetti di Big Data implementati e verificando se le loro esigenze sono supportate da un'offerta adeguata.

Il lavoro è stato condotto nel contesto della Ricerca 2015 condotta dall'Osservatorio Big Data Analytics & Business Intelligence della School of Management del Politecnico di Milano, il cui fine è mettere in risalto il valore strategico delle metodologie di Business Intelligence e dei Big Data Analytics per le imprese e la Pubblica Amministrazione e i potenziali vantaggi che possono essere ricavati.

L'elaborato parte dall'analisi della letteratura di riferimento sul tema Big Data (Capitolo 1), del quale abbiamo trattato diversi aspetti: le principali caratteristiche, i benefici che possono essere estratti dal loro sfruttamento, i freni che ne ostacolano l'adozione, le tecniche e le tecnologie di Big Data, i trend in atto, la governance dei sistemi, l'impatto nei principali settori e le azioni che devono essere messe in pratica per cogliere il potenziale offerto. Nel Capitolo 2 abbiamo approfondito l'effetti dei Big Data sul Marketing e sulla Customer Experience, considerando il compito sempre più difficile per le organizzazioni di soddisfare le esigenze dei consumatori, le quali possono cercare di far leva sull'enorme ammontare di dati che hanno a disposizione per riuscire in questa missione. Comprese le ragioni che spingono le aziende a sfruttare i Big Data nei più svariati ambiti e quindi a chiedere delle soluzioni abilitanti, abbiamo dedicato il Capitolo 3 all'analisi dell'offerta attuale in ambito *Advanced Analytics Platforms*, che rappresentano le soluzioni che consentono di estrarre maggiormente valore dai Big Data.

Abbiamo quindi presentato la metodologia di ricerca che illustra il percorso seguito nella realizzazione del presente elaborato (Capitolo 4). Dopo l'analisi della letteratura di riferimento e la definizione degli obiettivi, abbiamo rilevato i dati tramite survey online e interviste telefoniche, dalle quali abbiamo estratto i dati che ci hanno permesso di svolgere le analisi proposte e di stilare i casi di studio per un approfondimento delle iniziative in ambito Big Data.

Il Capitolo 5 rappresenta il fulcro di questo lavoro, in cui abbiamo cercato di rispondere alle domande di ricerca che ci siamo posti all'inizio relative al livello di maturità delle organizzazioni e delle funzioni aziendali, al sostegno garantito dall'offerta e agli ambiti progettuali delle funzioni di business risultate maggiormente mature. A tal fine ci siamo serviti di strumenti quali grafici, tabelle e matrici.

Abbiamo concluso il lavoro con una serie di interessanti casi di studio (Capitolo 6), preceduti dalla descrizione dei progetti Big Data implementati nei settori ai quali appartengono le aziende oggetto del caso.

## LETTERATURA DI RIFERIMENTO

La prima parte dell'analisi della letteratura di riferimento si propone l'obiettivo di descrivere il mondo dei Big Data, illustrando innanzitutto la definizione e le caratteristiche più importanti, le cosiddette 5V, ovvero il *volume*, la *velocità*, la *varietà*, la *veridicità* e il *valore*. L'analisi bibliografica ci ha permesso di individuare i benefici che i Big Data possono portare alle organizzazioni: la creazione di trasparenza, la scoperta dei bisogni dei consumatori, il miglioramento delle previsioni e delle prestazioni, la personalizzazione delle azioni, il supporto nel processo di decision making, la creazione di nuovi prodotti e servizi e innovativi modelli di business e l'incremento della produttività e della profittabilità delle aziende. Tuttavia perché le organizzazioni possano cogliere questi vantaggi, è necessario che superino una serie di barriere che abbiamo raggruppato in sei differenti categorie: *tecniche*, legate all'integrazione e alla qualità dei dati, *tecnologiche*, relative al livello di efficacia delle soluzioni adottate, *organizzative/gestionali*, riguardanti la gestione dei progetti Big Data all'interno delle funzioni, *legate alle competenze*, relative alla presenza di figure adeguate per queste iniziative, *culturali*, correlate all'apertura verso i Big Data, *economiche* e *legate alle privacy*. Abbiamo rivolto un'attenzione particolare all'assenza di specialisti di analisi dei dati, ovvero al Data Scientist e allo Chief Data Officer, che rappresenta l'ostacolo principale da superare, approfondendo il tema della Governance dei sistemi di Big Data Analytics (BDA). Abbiamo poi descritto le tecniche e le tecnologie abilitanti utilizzate per aggregare, manipolare, gestire e analizzare i Big Data e l'impatto di questi nel settore Sanitario, nella Pubblica Amministrazione, nel Manufacturing e nel Retail, mettendo in evidenza i benefici che possono essere conseguiti dal loro sfruttamento. Considerando l'enorme potenziale offerto dai Big Data e i freni che devono essere superati per poterli cogliere, abbiamo concluso riportando una serie di comportamenti e di azioni che le organizzazioni devono implementare a tal fine: focalizzarsi sulla gestione del cambiamento, ridefinire i ruoli e acquisire talenti, costruire una cultura orientata agli Analytics,

adottare una strategia orientata all'integrazione delle tecnologie, garantire la privacy e la sicurezza e fare in modo che CMO e CIO collaborino.

Nella seconda parte della letteratura ci siamo concentrati sull'impatto dei Big Data sul Marketing e sulla Customer Experience. I continui cambiamenti delle preferenze dei consumatori che occupano oggi una posizione di vantaggio rispetto alle aziende e quindi la necessità di garantire una personalizzazione one-to-one dei messaggi a loro rivolti, spinge le aziende a far leva sull'enorme ammontare di dati che hanno a disposizione e sui Marketing Analytics, approccio abilitato dalle tecnologie e supportato da modelli che sfrutta i dati di mercato e dei clienti per migliorare il processo di decision making in ambito Marketing. È per questo motivo che abbiamo deciso di dedicare il secondo capitolo a questo tema. Siamo partiti dal passaggio da un approccio classico al Marketing a uno predittivo, focalizzandoci quindi sui benefici apportati dai Big Data e dalle relative soluzioni, che consistono in un miglioramento della Customer Experience, il quale porta all'incremento del tasso di mantenimento e all'acquisizione di nuovi clienti, comportando un aumento della redditività delle aziende. Anche per sfruttare i *Marketing Analytics* devono essere superate diverse barriere che possiamo ancora raggruppare in tecniche, culturali, organizzative e legate alle competenze e seguiti una serie di passi che prevedono l'individuazione del miglior approccio analitico, l'integrazione delle competenze per la generazione di insight e la costruzione dell'approccio al centro dell'organizzazione. Ci siamo quindi focalizzati sui progetti Big Data in ambito Marketing, scendendo nel dettaglio di quelli di *Social Analytics*, che danno l'opportunità di valutare la credibilità del brand, identificare i bisogni dei clienti, caratterizzare il brand, ottimizzare la comunicazione, individuare gli influencer e monitorare i competitor. Abbiamo poi distinto cinque diversi livelli di sviluppo crescente dei Social Analytics, a seconda del grado di maturità analitica e tecnologica dei metodi utilizzati dalle organizzazioni: Social Unware, Social Monitoring, Social Listening, Social Intelligence e Social Integration. Gli altri progetti descritti sono il *Direct* e il *Digital Marketing*, la *Customer Micro-segmentation*, il *Location-based Marketing*, l'*In-store Analysis* e il *Cross-Selling/Up-Selling*. Abbiamo concluso approfondendo l'influenza dei Big Data sulla Customer Experience: le aziende devono far leva su di questi nel corso dell'intero viaggio decisionale del cliente per estrarre insight da integrare nei programmi di Marketing e anticipare i movimenti dei consumatori durante tutto il percorso con l'obiettivo di migliorare la loro esperienza.

Nell'ultimo capitolo di letteratura abbiamo studiato lo stato dell'offerta in ambito *Advanced Analytics Platforms*, i quali rappresentano gli strumenti che permettono di estrarre

maggiormente valore dai Big Data, prendendo come riferimento il *Magic Quadrant for Advanced Analytics* di *Gartner* aggiornato all'anno corrente. Le aziende stanno infatti maturando un'esigenza crescente di estrarre insight dall'analisi di enormi quantità di dati di tipologie differenti, andando oltre quella di query e reporting, garantita dalla più tradizionale Business Intelligence. Sul mercato il tasso di crescita degli Advance Analytics è infatti superiore a quello della più classica BI e questo è un trend che possiamo riscontrare anche in Italia. Con il Magic Quadrant i vendor sono stati suddivisi in quattro classi sulla base di due criteri: *l'ability to execute*, che si riferisce alla capacità nel rendere la loro vision una realtà di mercato che i clienti vedono come differenziata dalle altre e la *completeness of vision*, legata alla loro abilità nel comprendere come le forze dei mercati possano essere sfruttate per creare valore ai consumatori e a loro stessi. Le quattro categorie sono i *Leaders*, che hanno ottenuto un punteggio alto in entrambe le dimensioni, all'opposto i *Niche Players*, poi i *Challengers* che hanno un'ability to execute alta e una completeness of vision bassa e i *Visionaries*, che viceversa possiedono un'ability to execute bassa e una completeness of vision alta. Abbiamo concluso presentando come esempio l'offerta di quattro vendor appartenenti a quattro diverse classi.

## **OBIETTIVI DEL LAVORO**

I Big Data stanno cambiando il modo di raccogliere, analizzare e integrare i dati. Considerato l'incremento delle fonti informative disponibili, un numero sempre maggiore di aziende comprende i possibili vantaggi competitivi che i Big Data Analytics possono portare. Per comprendere effettivamente se le organizzazioni sono pronte ad adottare e a trarre i conseguenti vantaggi da queste soluzioni e quali sono in concreto gli approcci adottati dalle funzioni aziendali al loro interno siamo entrati nel merito delle medie e grandi organizzazioni italiane, analizzando questi temi. Il lavoro si propone primariamente di:

- valutare il livello di maturità alle soluzioni di Big Data Analytics delle medie e grandi organizzazioni italiane;
- verificare lo stato di maturità alle soluzioni di Big Data Analytics delle singole funzioni aziendali.

Perché queste abbiano concretamente la possibilità di sfruttare le potenzialità offerte da tali sistemi è necessario che l'offerta sia adeguata. Pertanto il terzo obiettivo della nostra analisi è:

- valutare il supporto dell'offerta di mercato.

Nello specifico abbiamo considerato l'offerta delle startup per l'alto tasso di innovazione che le contraddistingue, elemento fondamentale in un mondo come quello dei Big Data, all'avanguardia e che richiede l'introduzione di continue novità.

L'ultimo obiettivo che ci siamo posti, una volta determinate le funzioni aziendali più mature, è:

- indagare i loro ambiti progettuali.

In questo modo possiamo vedere in quale aree vengono effettivamente sfruttati i BDA e quali sono quelle dove sono previste le maggiori potenzialità nel futuro.

## **METODOLOGIA DI RICERCA**

La Ricerca ha coinvolto 160 C-level di varie funzioni aziendali e 91 CIO e Responsabili IT di medie e grandi organizzazioni italiane. Il percorso di ricerca che abbiamo seguito è costituito da più fasi sequenziali, la prima delle quali è l'analisi della letteratura di riferimento. In un secondo momento abbiamo definito gli obiettivi del lavoro e abbiamo proceduto con la definizione delle metodologie più idonee alla raccolta delle informazioni di rilievo in merito agli ambiti oggetto di analisi.

Abbiamo eseguito la rilevazione dei dati utilizzando tipologie di strumenti differenti per le diverse domande di ricerca. Per rispondere alle domande relative al livello di maturità delle organizzazioni e delle funzioni aziendali e agli ambiti progettuali delle funzioni più mature, ci siamo serviti di due *survey online*, una rivolta ai CIO e una ai C-level, e alle *interviste telefoniche*. La survey è stata realizzata con lo scopo di ottenere delle informazioni standard su cui fare delle valutazioni e individuare i casi aziendali più interessanti da approfondire attraverso le interviste. Queste ultime ci hanno consentito di indagare le singole realtà aziendali e fornire un quadro sui progetti di Big Data implementati nelle aziende appartenenti ai diversi settori.

Per rispondere alla terza domanda di ricerca, relativa al supporto dell'offerta ci siamo invece serviti del portale Crunchbase e in misura minore di altri siti, per effettuare il censimento di 498 startup internazionali e italiane nell'ambito Big Data Analytics. Nello specifico i dati rilevati sono quelli anagrafici, i finanziamenti complessivi ottenuti e i finanziamenti ricevuti dal 2012 e sulla base delle informazioni tratte abbiamo proceduto ad una classificazione per tipologia e origine dei dati analizzati e per ambiti applicativi.

Per trarre ulteriori informazioni e approfondire gli argomenti trattati abbiamo utilizzato i dati raccolti dai vari workshop relativi a diversi temi in ambito Big Data che si sono tenuti nel corso dell'anno dall'Osservatorio Big Data Analytics & Business Intelligence: il workshop *"I Big Data come leva strategica nel Retail"*, il workshop *"I Big Data nell'online business"* e il workshop *"I Big Data per la customer journey"*. Ulteriori spunti sono stati tratti dal *convegno conclusivo* tenuto dall'Osservatorio Big Data Analytics & Business Intelligence.

Dopo aver raccolto le informazioni d'interesse, abbiamo svolto l'analisi dei dati e la redazione dei casi di studio per perseguire gli obiettivi della ricerca.

## RISULTATI OTTENUTI

L'analisi dei risultati ci ha permesso di rispondere alle 4 domande di ricerca che siamo posti. Relativamente al livello di maturità delle organizzazioni è emerso che se da un lato ce ne sono alcune che stanno cercando di far leva sui Big Data a 360°, conscie delle opportunità strategiche che questi offrono, dall'altro ce ne sono alcune maggiormente sospettose, che risultano quindi frenate. Complessivamente possiamo dire che nel percorso che porterà all'adozione dei progetti di Big Data Analytics siamo ancora in uno stadio iniziale. In particolare le organizzazioni sono state classificate in quattro categorie: *pionieri*, che riconoscono l'importanza dei Big Data per la strategia aziendale e sono in una fase avanzata di implementazione degli Analytics; *reattivi*, che sviluppano un piano condiviso per la gestione dei Big Data essendo consapevoli dei vantaggi che questi possono apportare, ma sono in ritardo nell'adozione delle soluzioni; *tradizionalisti*, i quali non hanno una mentalità aperta verso le possibilità che i Big Data mettono a disposizione all'intera organizzazione, tuttavia implementano dei sistemi in specifici ambiti; *scettici*, che non hanno ancora compreso a pieno le opportunità offerte dai Big Data e per questo motivo non hanno molto interesse ad introdurre strumenti di Big Data Analytics.

Relativamente alla seconda domanda di ricerca, è emerso che è la funzione di *Marketing e Vendite* a presentare il più alto livello di maturità nei Big Data Analytics e ad essere quindi pronta a cogliere le grandi opportunità offerte da queste soluzioni, così come l'*Amministrazione, Finanza e Controllo* che mostra un livello medio/alto. Gli *Acquisti* e la *Logistica e Produzione* si trovano invece in uno stato intermedio, la *Ricerca e Sviluppo* in un medio-basso, mentre le *Risorse Umane* risultano essere la funzione più arretrata in quest'ambito.

Scendendo nel dettaglio, la funzione di Marketing e Vendite è risultata avere un maggior livello di maturità nelle tre dimensioni indagate relative alla conoscenza e allo stato attuale e futuro di diffusione dei BDA, ai dati analizzati e alle soluzioni adottate. In particolare è l'unica che *mixa* adeguatamente dati interni ed esterni, soprattutto quelli social, evidenziando l'importanza del fenomeno per la definizione delle strategie di Marketing, e che ha a disposizione soluzioni che garantiscono un supporto efficace ed adeguato alle esigenze.

Un aspetto critico che interessa tutte le funzioni, compreso il Marketing e Vendite, è l'utilizzo da parte dei sistemi di Analytics prevalentemente di dati "tradizionali" (dati di natura strutturata e interni): è possibile confermare tale risultato dalle informazioni delle survey condotte, in quanto per la grande maggioranza vengono analizzati dati transazionali, record e documentazioni office, quindi di natura strutturata. Altri dati analizzati appartenenti a questa categoria, ma in misura nettamente inferiore, sono i log Data, i dati di localizzazione e GPS, quelli M2M generati dai sensori e quelli originati da banche dati open.

Il fatto che le varie funzioni aziendali, chi più chi meno, non riescono ancora a cogliere al 100% le opportunità offerte dai sistemi di BDA, può essere giustificato dalla serie di ostacoli che si trovano ad affrontare, condivisi da tutte. Il principale freno è rappresentato dalla mancanza di figure specializzate quali il Data Scientist e lo Chief Data Officer; altre barriere importanti sono la difficoltà di integrazione con altri applicativi, la complessità nello stimare il ritorno dell'investimento, gli elevati tempi di implementazione e la mancanza di coinvolgimento del top management e l'ingente investimento.

L'analisi è quindi proseguita con l'indagine sul ruolo delle startup e sul supporto fornito da loro in termini di offerta. Innanzitutto abbiamo considerato le startup che hanno ricevuto finanziamenti da investitori istituzionali dal 2012 ad oggi, per un totale di 14,48 miliardi di dollari. Esse offrono per la maggior parte soluzioni orientate a diverse industry che analizzano sia dati strutturati che dati non strutturati e sia dati esterni che dati interni. I risultati più interessanti sono derivati dalla classificazione per ambiti applicativi, per cui sono state identificate 3 categorie: *Infrastructure*, a sua volta suddivisa in Database, Technologies, Storage e Cross-infrastructure, *Analytics* per la quale abbiamo identificato le classi di Data Visualization, Search, Location/people/events, Analytics platform e Social Analytics e *Application*, suddivise negli ambiti di Marketing&Sales, CRM & Customer Experience, Finance, Supply Chain, Logistica & Operations, IT, Human Resource e Security, per un totale quindi di 16 categorie. Le soluzioni principalmente offerte sono le Technologies, gli Analytics platform e i Social Analytics e le Application di Marketing & Sales, CRM & Customer Experience.



Per verificare il supporto dell'offerta alle varie funzioni aziendali abbiamo identificato a quali di queste le soluzioni offerte dalle startup appena descritte si rivolgono ed è emerso che la funzione caratterizzata dall'offerta più ampia è il Marketing e Vendite.

Abbiamo quindi analizzato il livello della domanda, identificando quante soluzioni diverse vengono richieste dalle funzioni aziendali ed è risultato che è ancora il Marketing e Vendite ad essere caratterizzato dal livello di domanda più alto, in linea con il suo alto livello di maturità.

Infine abbiamo mappato su una matrice (che riportiamo anticipatamente data la sua importanza) il livello dell'offerta e della domanda, che abbiamo valutato considerando diverse proxy, per individuare il posizionamento delle funzioni. I principali risultati sono l'ottima posizione occupata dalla funzione *Marketing e Vendite*, caratterizzata da un alto livello di domanda al quale corrisponde un'offerta adeguata, a testimonianza di come le startup abbiano compreso le esigenze delle aziende, la situazione critica per *Ricerca e Sviluppo*, dove sia la domanda che l'offerta di Big Data Analytics sono scarse e la posizione particolare quella occupata dalle *Risorse Umane*, dove possiamo osservare un grande traino dell'offerta che propone soluzioni per ottimizzare e migliorare i vari processi, ma che non vede uno stesso riscontro nella domanda.

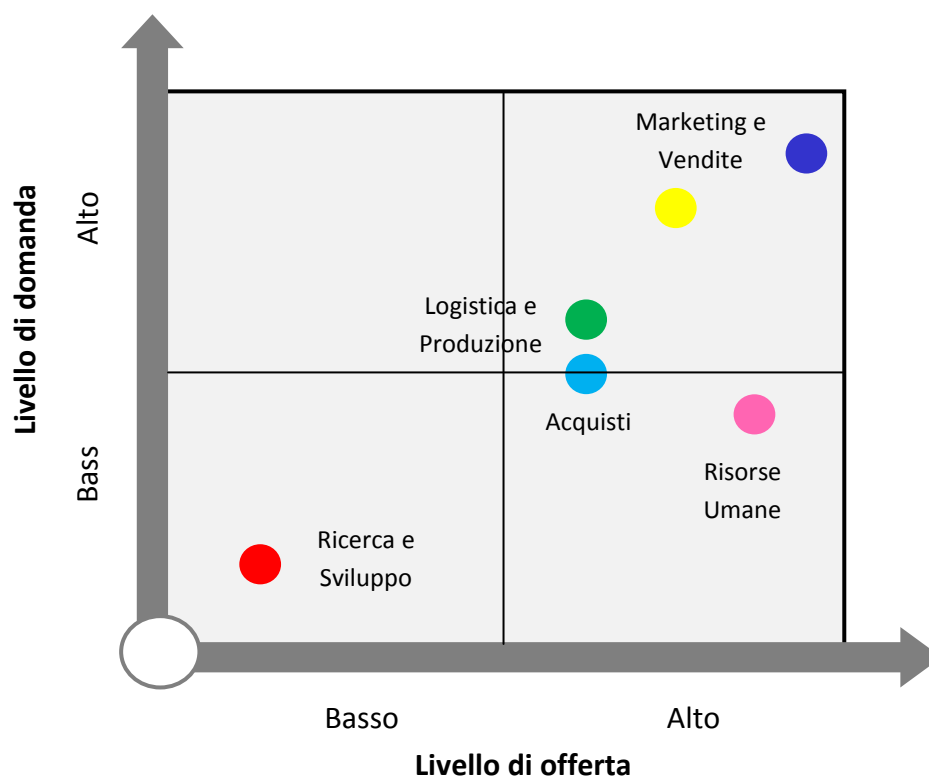


Grafico 1: Posizionamento delle funzioni aziendali rispetto nella matrice livello di domanda/livello di offerta

Infine per quanto riguarda gli ambiti progettuali la situazione mostrata è molto promettente, considerando le iniziative previste nel prossimo futuro, che mirano a far leva sempre di più sui Big Data, anche se attualmente si trovano in uno stadio iniziale. Nella funzione di Marketing e Vendite abbiamo infatti individuato alcuni progetti già ampiamente presenti che sfruttano i Big Data, quali quelli di Direct Marketing, di Cross e Up-Selling e di Social Analytics e il grande interesse a sfruttare l'enorme mole di dati a disposizione, tra cui quelli di geolocalizzazione, quelli social e quelli di navigazione sul web per altre iniziative.

## CONCLUSIONI

Cerchiamo ora di sottolineare brevemente i punti chiave e le conclusioni tratte della nostra analisi.

Innanzitutto la maggior parte delle organizzazioni ha compreso l'importanza di estrarre insight dai dati e sta iniziando a far leva sui Big Data, ma risulta ancora lontana dall'adottare una strategia di business Data-driven. Inoltre c'è sempre una parte di "scettici", che non sono ancora riusciti a cogliere le opportunità da loro offerte.

Considerando le singole funzioni aziendali, è emerso che il Marketing è la funzione più matura nell'ambito dei Big DataAnalytics e con le maggiori prospettive di crescita. Nelle diverse dimensioni indagate, relative allo stato di diffusione dei Big DataAnalytics, della consapevolezza dell'effetto strategico dei Big Data, dei dati utilizzati e delle soluzioni adottate si trova in una condizione migliore rispetto a tutte le altre funzioni, portandola ad occupare una posizione privilegiata. Solo la funzione di Amministrazione, Finanza e Controllo sembra tenere il passo, gli Acquisti e la Logistica e Produzione si trovano in uno stadio intermedio, mentre la Ricerca e Sviluppo e le Risorse Umane sono ancora in una fase arretrata. Considerando quindi la situazione nel complesso possiamo individuare alcuni elementi motivo di soddisfazione e altri meno: se da un lato il lavoro da fare per le funzioni, compreso il Marketing e Vendite, per raggiungere un maggior livello di maturità è ancora lungo, dall'altro possiamo osservare l'impegno di tutte per migliorare.

Le esigenze delle funzioni aziendali sono ben coperte dall'offerta che rappresenta l'area nella quale abbiamo individuato il maggior riscontro positivo. Anche in questo ambito è il Marketing e Vendite a trovarsi nella posizione migliore, considerando l'ingente quantità di soluzioni che gli viene proposta. Le 498 startup censite mettono a disposizione soluzioni che vanno dalle Infrastrutture, agli Analytics fino ad Applicazioni rivolte a uno o più ambiti applicativi e la loro

offerta in un mercato così innovativo come quello dei Big Data, è essenziale dato il loro alto livello di innovazione.

Anche l'analisi degli ambiti progettuali delle funzioni più mature mostra una situazione molto promettente, considerando le iniziative previste nel prossimo futuro, che puntano a sfruttare sempre di più i Big Data.

## ABSTRACT

Questo lavoro di tesina parte da un'analisi della letteratura sul tema Big Data, del quale abbiamo esplorato le principali caratteristiche, i vantaggi che possono essere ricavati dal loro sfruttamento, gli ostacoli che impediscono il loro utilizzo, la governance di tali sistemi e lo stato dell'offerta.

Dall'analisi bibliografica sono emerse le enormi opportunità che i Big Data offrono alle organizzazioni e a tutte le sue funzioni aziendali e questo è un aspetto che abbiamo voluto approfondire analizzando la situazione italiana all'interno delle medie e grandi organizzazioni.

Nello specifico siamo partiti con l'indagine sullo stato di maturità dei Big Data Analytics, determinando come queste si trovino ancora all'inizio del percorso che porterà alla loro piena adozione. Se da un lato infatti alcune hanno già integrato tali soluzioni nei loro piani strategici, conscie delle potenzialità offerte, dall'altro lato una buona parte mostra ancora un certo scetticismo. Abbiamo riscontrato una situazione analoga all'interno delle funzioni aziendali, trovandone alcune, tra tutte il Marketing e Vendite, pronte a sfruttare le soluzioni di BDA e altre in una fase ancora arretrata. Per la valutazione abbiamo analizzato una serie di dimensioni, identificandone alcune più mature, come la qualità nella gestione dei dati a disposizione e altre in cui il lavoro da fare è ancora molto, quali l'utilizzo di dati di natura non strutturata e di origine esterna, sempre più importanti considerando l'esplosione del fenomeno social.

Il lavoro è proseguito con un'indagine sulle startup che offrono soluzioni in quest'ambito al fine di indagare il supporto fornito alle funzioni di business. 498 startup internazionali e italiane sono state censite e classificate secondo alcune dimensioni, identificando come queste riescano a coprire molto bene le esigenze di tutte, specialmente dell'unità di Marketing&Vendite.

Abbiamo quindi proseguito la Ricerca approfondendo gli ambiti progettuali delle funzioni aziendali più mature, identificando iniziative già presenti e ad un buon punto ed altre molto promettenti per il futuro, che mettono in evidenza la volontà di far leva in misura sempre maggiore sui Big Data.

## INTRODUZIONE

*“So what’s getting ubiquitous and cheap? Data. And what is complementary to Data? Analysis. So my recommendation is to take lots of courses about how to manipulate and analyze Data: Databases, machine learning, econometrics, statistics, visualization, and so on.”*

Quest’affermazione di Hal Varian, Chief Economist di Google e noto professore dell’Università di Berkeley in California, mette in evidenza come i dati rappresentino una fonte molto importante che le organizzazioni devono analizzare per sfruttare a loro vantaggio.

Al giorno d’oggi stiamo assistendo ad una vera e propria esplosione di dati: le aziende catturano miliardi di miliardi di bytes di informazioni relative a clienti, fornitori e operazioni e generano un volume crescente di dati transazionali; milioni di networked sensors sono incorporati in device quali cellulari, macchine industriali e automobili che rilevano, creano e comunicano i dati; milioni di persone in un mondo ormai digitalizzato comunicano, acquistano, condividono, fanno ricerche tramite social media, smarhpone e PC, contribuendo alla creazione di questa enorme massa di dati. I Big Data, che sono ormai parte di molti settori e di molte funzioni dell’economia globale, rappresentano un driver fondamentale per l’innovazione e la crescita e una fonte significativa per la creazione di valore e vantaggio competitivo, tanto da comportare una rivoluzione del management e da suscitare entusiasmo in tutto il mondo.

L’era dei Big Data è quindi una rivoluzione che cambia il modo di essere e di fare di un’impresa. È quindi necessario per le aziende attrezzarsi in modo da riuscire a sfruttarli a pieno, portando anche numerosi vantaggi ai consumatori finali.

# 1 L'ERA DEI BIG DATA

## 1.1 COSA SONO I BIG DATA?

Prima di iniziare la nostra Ricerca, la domanda che dobbiamo porci è: "Cosa sono esattamente i Big Data?". Essi si riferiscono a un dataset la cui dimensione va al di là della capacità di un database normale di catturare, memorizzare, gestire e analizzare i dati (Manyika J., 2011). Si tratta di una definizione soggettiva in cui non viene esplicitata la dimensione che deve avere un Dataset per poter essere definito Big Data, in quanto essa aumenterà sicuramente nel tempo dati i continui avanzamenti tecnologici. Le dimensioni variano nei diversi settori, da dozzine di terabyte a centinaia di petabyte (1000 terabyte), in base anche agli svariati strumenti software a disposizione.

In questa definizione emerge una delle cosiddette 5V che caratterizzano i Big Data, ovvero il volume, le altre sono velocità, varietà, veridicità e valore (Opresnik D., 2015). Il *volume* fa appunto riferimento all'enorme massa di dati generata attraverso numerosi canali: ogni giorno Google elabora circa 24 petabytes di dati, un motore a reazione può generarne 10 terabyte in 30 minuti così come i contatori intelligenti e i macchinari per l'industria pesante tra cui le raffinerie di petrolio e gli impianti di perforazione. La *velocità* si riferisce alla rapidità con cui i dati vengono acquisiti e utilizzati, in aumento date transazioni sempre più frequenti e veloci: le aziende non solo raccolgono i dati più velocemente, ma cercano di sfruttarli il prima possibile, spesso in real time. La *varietà* è legata alle differenti tipologie di dati disponibili provenienti da un numero crescente di fonti di dati sia strutturati sia non strutturati; in particolare è possibile identificare cinque categorie di informazioni che costituiscono i Big Data:

- dati generati da smartphone e altri dispositivi mobile relativi a persone, attività e localizzazione, tra cui dati RFID (radio-frequency identification), dispositivi che tracciano il prodotto, e dati da dispositivi di controllo come i contatori per il monitoraggio dell'acqua o del gas;
- dati di vendita e pricing, dati generati dall'attività delle carte fedeltà e degli eventi promozionali;
- computer log Data, come i clickstreams dai siti web;
- informazioni dai social media come Twitter e Facebook;
- social multimediali e altre informazioni da Flickr, YouTube e siti simili.

La *veridicità* riguarda la questione relativa alla qualità dei dati e al loro livello di sicurezza, la cui garanzia rappresenta una sfida molto importante. Il *valore* è un aspetto fondamentale: per poter sfruttare i Big Data è necessario saper agire per poter estrarne valore e quindi incrementare la produttività e la competitività delle aziende e creare surplus economico per i consumatori.

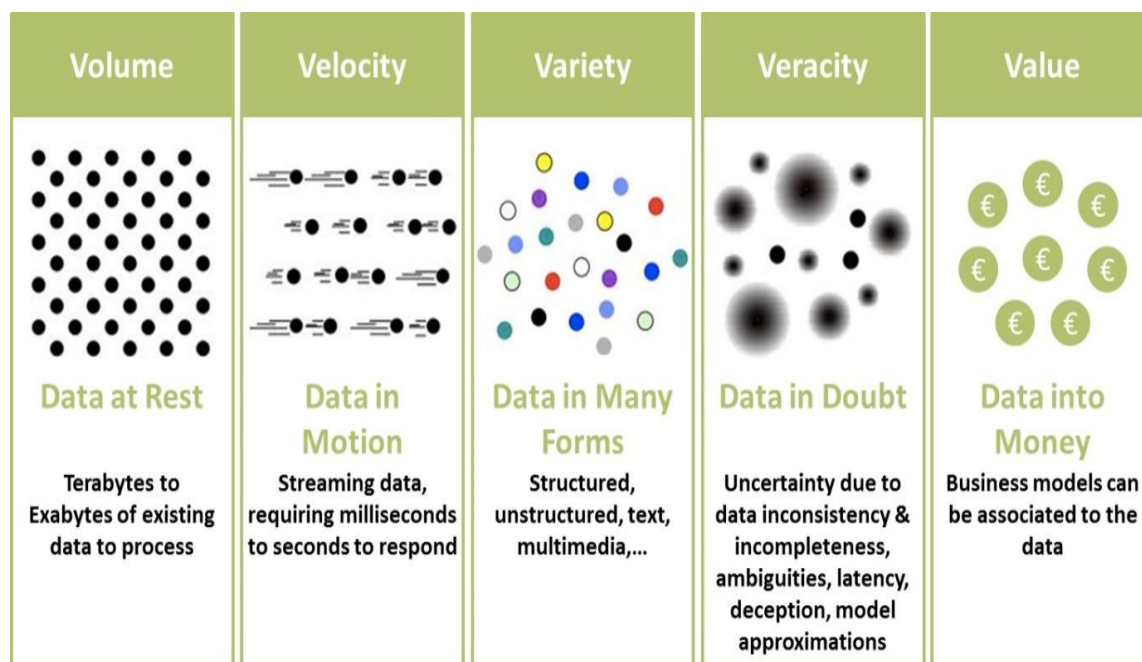


Figura 1: Le 5V dei Big Data (Google, 2014)

La capacità di memorizzare e aggregare i dati e quindi di utilizzare i risultati per svolgere analisi profonde migliora continuamente grazie alla disponibilità di strumenti software e tecniche sempre più sofisticate combinate a una crescente potenza di calcolo. Stiamo assistendo anche ad un enorme cambiamento della capacità di generare, comunicare, condividere e accedere ai dati dovuto all'aumento del numero di persone, strumenti e sensori ora connessi da reti digitali. Per capire la grandezza del fenomeno, basta osservare la figura nella pagina successiva che mostra quanti dati vengono generati in un minuto (Osservatorio Big Data Analytics & Business Intelligence, maggio 2015).

## 1.2 BENEFICI CONSEGUIBILI DALL'UTILIZZO DEI BIG DATA

I Big Data rappresentano una grande opportunità per le aziende e per le economie nazionali in quanto consentono di ottenere benefici significativi, che elenchiamo di seguito.

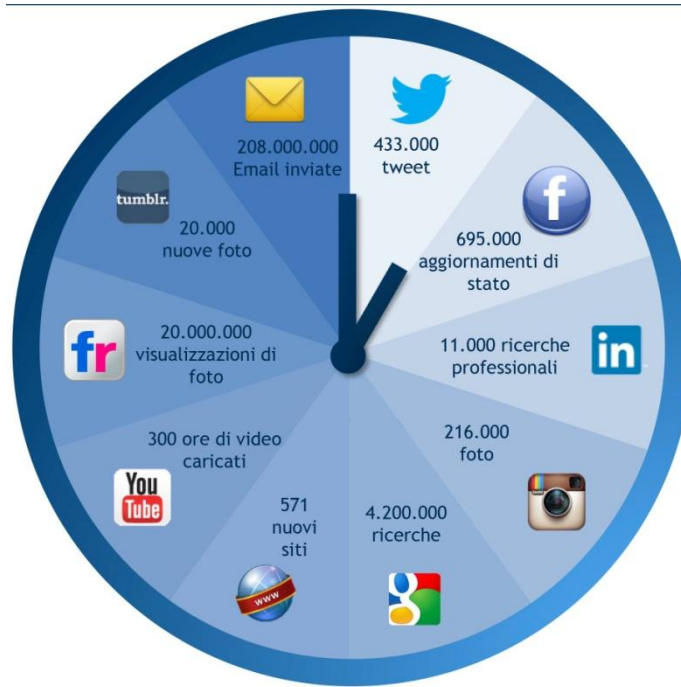


Figura 2: Dati generati in un minuto (Osservatorio Big Data Analytics & Business Intelligence, maggio 2015)

- ✓ **Creare trasparenza.** Un accesso facile e tempestivo ai Big Data rende disponibile una maggiore quantità di informazione e facilita la condivisione dei dati tra le diverse unità organizzative di un'impresa (Bernice P., 2013). Per esempio i dati delle unità di R&S, produzione e ingegneria di un'azienda possono essere integrati al fine di favorire il concurrent engineering, tagliando i tempi e migliorando la qualità (Manyika J., 2011).
- ✓ **Scoprire i comportamenti nascosti e i bisogni dei consumatori.** La disponibilità quasi in real time di dati da smartphone fornisce caratteristiche dettagliate sui clienti e sul loro complesso processo decisionale quando fanno acquisti: i Big Data permettono infatti di identificare i modelli comportamentali dei consumatori e far luce sulle loro intenzioni (Michael K., 2013).
- ✓ **Rivelare le variabilità delle performance e migliorare le prestazioni.** La creazione e la memorizzazione di dati transazionali in forma digitale consente poi alle aziende di avere dati più accurati e dettagliati su svariate performance, dallo stato dei magazzini ai giorni di malattia del personale, in tempo reale o quasi. Inoltre esse, utilizzando i dati per analizzare la variabilità delle prestazioni e per capirne le cause più profonde, possono ottenere risultati migliori (Manyika J., 2011).



Una nota azienda italiana che offre soluzione per la gestione dei dati aziendali, su richiesta di una compagnia assicurativa, desiderosa di migliorare le sue performance, ha analizzato le vendite del suo canale call center, identificato i pattern di successo, determinando quindi la telefonata perfetta.

- ✓ **Personalizzare le azioni.** I Big Data consentono di creare specifici segmenti di clienti e di personalizzare prodotti e servizi sulla base delle loro esigenze. Si tratta di un grande beneficio per vari settori: le aziende di beni di consumo per esempio stanno iniziando ad utilizzare tecniche di Big Data per realizzare promozioni e pubblicità personalizzate per i diversi cluster (Manyika J., 2011).
  
- ✓ **Migliorare le previsioni.** L'utilizzo dei Big Data e di tecniche adeguate per il loro sfruttamento portano a migliori previsioni e migliori previsioni fruttano migliori decisioni. Per esempio, le principali compagnie aeree statunitensi, venute a conoscenza del disallineamento tra l'orario previsto e quello effettivo di atterraggio, hanno deciso di utilizzare il servizio RightEta offerto da PASSUR Aerospace, un fornitore di tecnologie di supporto decisionale nel campo dell'aviazione, per risolvere questo problema. Più di 155 installazioni raccolgono un ampio range di informazioni su tutti i voli che vedono, generando un costante flusso di dati digitali che vengono analizzati, quindi RightEta si chiede che cosa sia successo in passato quando un dato aereo si è avvicinato ad un dato aeroporto in determinate condizioni e quando effettivamente sia atterrato. In questo modo la compagnia aerea è riuscita ad eliminare il divario tra arrivo previsto e arrivo effettivo dei voli e questo miglioramento delle previsioni ha portato ad un valore di 7 milioni di dollari all'anno in ciascun aeroporto (McAfee A., 2012).  
Diverse Banche italiane stanno mettendo in atto un progetti finalizzati alla raccolta di dati social per arricchire le informazioni sui loro clienti e sfruttarle per predire il tasso di churn tramite adeguati modelli.
  
- ✓ **Supportare le persone nel processo di decision making.** Utilizzando Analytics sofisticati su interi Dataset è possibile automatizzare e migliorare i processi decisionali, minimizzare i rischi e scoprire preziosi insight, benefici che non possono essere perseguiti con l'analisi e la gestione di piccoli campioni di dati tramite i fogli di calcolo. I rivenditori per esempio possono utilizzare algoritmi che consentono la messa a punto automatica e l'ottimizzazione degli inventari e dei prezzi a partire dai dati in tempo reale relativi alle vendite nei negozi e a quelle online (Manyika J., 2011).

- ✓ **Creare nuovi prodotti e servizi, nuove tipologie di aziende e innovativi modelli di business.** Le società possono sfruttare i Big Data per realizzare nuovi prodotti e nuovi servizi: molte imprese manifatturiere per esempio stanno utilizzando i dati relativi all'utilizzo di prodotti attuali per migliorare lo sviluppo di modelli futuri e per creare servizi post-vendita innovativi; la disponibilità in real time di dati relativi alla location sta comportando lo sviluppo di nuovi servizi che si servono di questi dati, come le assicurazioni danni basate su dove e come le persone guidano le loro automobili. Nasceranno inoltre aziende che si occuperanno di aggregare ed analizzare i dati aziendali relativi a prodotti, servizi, fornitori, consumatori e loro preferenze e si assisterà addirittura a nuovi modelli di business (Manyika J., 2011).
  
  - ✓ **Incrementare la produttività e la profittabilità delle aziende.** Lo sfruttamento dei Big Data può portare ad un aumento dell'efficacia e dell'efficienza delle imprese, le quali potranno realizzare più output utilizzando meno input e migliorare il livello di qualità dell'output stesso. Questo vantaggio interesserà svariati settori: in quello manifatturiero è stata prevista una riduzione di più del 50% dei costi di sviluppo del prodotto e di quelli di assemblaggio e una diminuzione di più del 7% del capitale circolante, nel Retail in America è stato stimato un incremento del margine operativo netto del 60% e un aumento della crescita di produttività annua pari allo 0,5-1%, mentre nel settore sanitario americano e nel settore della PA europeo è stato previsto che la produttività aumenterà rispettivamente dello 0,7% e dello 0,5% annuo (Manyika J., 2011).
- Uno studio condotto al MIT Center for Digital Business ha dimostrato che le aziende Data-driven perseguono migliori performance finanziarie e operative. 330 dirigenti di aziende pubbliche del Nord America sono state intervistate su aspetti quali la gestione organizzativa e tecnologica e sono stati raccolti i risultati dai loro report annuali e da fonti indipendenti: le tre migliori aziende di ciascun settore nell'utilizzo dei Big Data, in media, sono più produttive dei loro competitor per il 5% e più profittevoli per il 6% (McAfee A., 2012).

Questo elenco di benefici mette in evidenza come l'investimento nei Big Data porti alla creazione di valore per le aziende e quindi all'ottenimento di vantaggio competitivo nel lungo termine. Risulta quindi fondamentale per loro sviluppare competenze in questo ambito, pena il declino in un mondo Big Data.

### 1.3 BARRIERE ALL'UTILIZZO DEI BIG DATA

Nonostante le opportunità offerte dai Big Data siano enormi, c'è ancora un certo scetticismo all'interno delle aziende sui reali benefici apportati a causa degli scarsi risultati ottenuti in pratica. In uno studio condotto di recente (fine 2014) dal McKinsey Global Institute (MGI) rivolto agli Analytics leader di alcune importanti aziende americane impegnate nella realizzazione di progetti di Big Data e di advanced Analytics, è emerso come l'utilizzo di queste tecniche abbia portato ad un aumento dei ricavi e ad un abbassamento dei costi inferiore all'1% per i tre quarti degli intervistati (Court D., 2015).

Esistono quindi una serie di barriere all'utilizzo dei Big Data da considerare, che possono essere classificate in 6 categorie: barriere tecniche, barriere legate alle competenze, barriere organizzative/gestionali, barriere culturali, barriere economiche e barriere legate alla privacy.

Tipologie barriere		Descrizione
Barriere tecniche	Difficoltà di integrazione dei dati	Integrare dati disomogenei, provenienti da svariate fonti strutturate e non, è un compito molto complesso, così come integrare dati che arrivano da diverse aree aziendali. Di conseguenza estrarre <i>insight</i> e trasformare questi in azioni non è così semplice.
	Basso grado di influenza del business	Molto spesso gli Analytics non portano a risultati significativi per le aziende, frenando ulteriormente la loro adozione da parte delle organizzazioni. Il dirigente di una famosa casa automobilistica ha investito recentemente in un'iniziativa per capire come i social media possano essere utilizzati per migliorare la

		pianificazione della produzione e le previsioni e ha osservato come l'analisi abbia fatto emergere dettagli interessanti sulle preferenze dei consumatori, ma non abbia fornito una guida su come migliorare effettivamente l'approccio alle previsioni (Court D., 2015).
	Scarsa qualità dei dati	Non sempre le aziende hanno a disposizione dati aggiornati e affidabili e questo è un altro ostacolo da superare.
Barriere legate alle competenze	Difficoltà di comprensione degli strumenti analitici e di quantificazione dei benefici	I manager di linea e i vari utilizzatori non capiscono gli strumenti analitici utilizzati o i consigli che suggeriscono, soprattutto quando questi sono difficili da utilizzare o quando non sono incorporati con i flussi di lavoro e i processi. Risulta perciò complesso per loro anche stimare i benefici derivanti dall'utilizzo dei tool e quindi il ritorno sull'investimento.
	Carenza di talenti	La scarsità di talenti, ovvero di persone con competenze di statistica e di machine learning e manager in grado di sfruttare gli insight dai Big Data nei loro business, rappresenta una delle barriere più importanti. All'interno delle aziende mancano quindi figure organizzative specializzate, quali il Data Scientist e lo Chief Data Officer, che

		<p>approfondiremo più avanti. Negli Stati Uniti nel 2018 si stima una domanda di talenti pari 440000/490000, un'offerta pari a 300000 e quindi un gap di 140000/190000. Persone con questo tipo di abilità sono difficili da trovare in quanto lo sviluppo di tali competenze richiede anni di formazione e inoltre non si può pensare di colmare questo gap cambiando i requisiti di laurea o aspettando persone che si laureeranno in questo ambito o importando talenti, ma risulta necessario riqualificare individui che sono già a disposizione. (Manyika J., 2011).</p>
	Difficoltà nella scelta del tool adatto	<p>La grande varietà di tool, che cambiano molto velocemente e fanno cose molto diverse, rende il processo di scelta molto difficile per le aziende, dato il livello non ancora adeguato di competenze nell'ambito Big Data.</p>
Barriere organizzative/gestionali	Mancanza di commitment da parte del top management	<p>I top manager non sono coinvolti nelle iniziative di Big Data, verso le quali mostrano poco interesse e quindi non forniscono l'aiuto necessario.</p>
Barriere culturali	Inerzia	<p>La maggior parte delle aziende non è ancora pronta e del tutto aperta alle innovazioni che i Big Data potrebbero portare, in quanto il loro sfruttamento richiederebbe significativi cambiamenti culturali e organizzativi. Per esempio,</p>

		<p>per un'azienda sarebbe importante avere a disposizione dati in tempo reale e meccanismi automatici di pricing, ma se i processi di management prevedono di stabilire i prezzi su base settimanale, l'organizzazione non sarà in grado di sfruttare le opportunità offerte dalla tecnologia (Court D., 2015). Il problema è che nelle organizzazioni non c'è consapevolezza sull'impatto che i Big Data avranno sulla gestione aziendale. Questo comporta lunghi tempi di implementazione.</p>
Barriere economiche	Investimento elevato	<p>Le iniziative Big Data richiedono ingenti spese in termini di tecnologie implementate e di nuove figure professionali da assumere.</p>
Barriere legate alle privacy		<p>I consumatori non vogliono che le loro informazioni personali, come i personal location Data e i dati elettronici generati dal loro uso di Internet, vengano utilizzate dalle aziende, soprattutto perchè non sanno dove e come queste verranno sfruttate dalle organizzazioni, le quali devono considerare anche le leggi relative alle privacy dei diversi Paesi.</p> <p>Tools che consentono di tracciare ogni movimento dei dipendenti e di misurare continuamente le loro</p>

	performance fanno gli interessi delle organizzazioni e non dei singoli individui, che vedono minacciata la loro privacy. Le aziende devono quindi preservare la privacy individuale e questo le limita nello sfruttamento dei Big Data.
--	---

Tabella 1: Barriere all'utilizzo dei Big Data

#### 1.4 TECNICHE E TECNOLOGIE PER L'ANALISI DEI BIG DATA

Di seguito riportiamo in una tabella le principali tecniche e le tecnologie utilizzate per aggregare, manipolare, gestire e analizzare i Big Data.

Tecniche	Tecnologie
<i>A/B testing</i> : tecnica in cui un gruppo di controllo viene confrontato con gruppi di test al fine di determinare quali modifiche e azioni miglioreranno una data variabile obiettivo, come il tasso di risposta a una campagna di Marketing.	<i>Cassandra</i> : sistema open source di gestione dei Database, progettato per trattare grandi quantità di dati su un sistema distribuito. Questo sistema è stato sviluppato originariamente da Facebook e ora è gestito come progetto dalla fondazione Apache Software.
<i>Classificazione</i> : insieme di tecniche che permettono di identificare a quali categorie appartengono nuovi dati, basandosi su un training set i cui i dati sono già stati categorizzati.	<i>Database NewSQL</i> : classe di moderni sistemi di Database relazionali che cercano di fornire le stesse prestazioni scalabili dei sistemi NoSQL per l'elaborazione delle transazioni online in lettura e scrittura
<i>Cluster analysis</i> : metodo statistico per classificare gli oggetti, che divide un grande	<i>Database relazionali</i> : Database costituito da un insieme di tabelle, in cui dati sono

gruppo in piccoli gruppi caratterizzati internamente da omogeneità non nota in anticipo.	memorizzati in righe e colonne. I sistemi di gestione dei Database relazionali memorizzano solo dati strutturati e utilizzano per lo più il linguaggio SQL (linguaggio progettato per la gestione dei dati di questi Database. Questa tecnica permette di inserire, interrogare, aggiornare e cancellare i dati, di controllarne l'accesso e di gestire la struttura del Database).
<i>Crowdsourcing</i> : tecnica utilizzata per raccogliere dati, sottoposta a un grande gruppo di persone o a una comunità, attraverso per esempio i network media come il Web.	<i>Database non relazionali (NoSQL)</i> : Database che non memorizza i dati in tabelle (righe e colonne).
<i>Data fusion e Data integration</i> : insieme di tecniche che integrano e analizzano dati provenienti da diverse fonti al fine di sviluppare insight più efficienti e accurati rispetto a quelli ottenuti esaminando una singola fonte.	<i>Data warehouse</i> : Database specializzato per il reporting e spesso utilizzato per memorizzare grandi quantità di dati strutturati. I dati sono caricati attraverso strumenti di ETL (extract, transform e load) da Database operazionali e i report sono generati tramite strumenti di BI.
<i>Data mining</i> : insieme di tecniche di classificazione, cluster analysis, regole associative e regressione, che permette di estrarre modelli da grandi dataset combinando metodi statistici e di machine learning con la gestione dei database.	<i>Google File System</i> : sistema proprietario di file distribuito, sviluppato da Google. È stata l'ispirazione per Hadoop.
<i>Machine Learning</i> : parte della computer science riguardante la progettazione e lo sviluppo di algoritmi che consentono ai	<i>Hadoop</i> : software framework open source che processa grandi Data set relativamente a un certo tipo di problema su un sistema



computer di identificare i comportamenti basandosi su dati empirici e in particolare di riconoscere schemi complessi e prendere decisioni intelligenti.	distribuito. Il suo sviluppo è stato ispirato da Google File System e da MapReduce di Google. Hadoop è stato sviluppato originariamente da Yahoo! E ora è gestito come progetto di Apache Software Foundation.
<i>Modelli predittivi</i> : tecniche in cui viene creato o scelto un modello matematico per prevedere la probabilità di un risultato.	<i>HBase</i> : Database open source, distribuito e non relazionale, modellato sul BigTable di Google. È stato sviluppato originariamente da Powerset e ora è gestito come progetto di Apache Software Foundation, come parte di Hadoop.
<i>Natural language processing (NLP)</i> : insieme di tecniche di computer science e linguistica che si ricorrono ai computer per analizzare il linguaggio umano.	<i>Hive</i> : software sviluppato da Facebook e da poco open source che gestisce e interroga dati memorizzati in un Hadoop cluster utilizzando un linguaggio simile al linguaggio SQL. È più familiare rispetto ad Hadoop per gli utilizzatori di strumenti di BI.
<i>Network analysis</i> : insieme di tecniche utilizzate per caratterizzare le relazioni tra nodi in un grafo o in una rete. Nella social network analysis vengono analizzate le relazioni tra individui di una comunità o di un'organizzazione, per esempio come viaggiano le informazioni.	<i>In-memory Database</i> : Database Management System che gestisce i dati nella memoria centrale, molto più veloce dei DBMS su memorie di massa, ma le moli di dati sono molto inferiori.
<i>Ottimizzazione</i> : insieme di tecniche numeriche utilizzate per riprogettare sistemi complessi e processi al fine di migliorare le performance relativamente a uno o più aspetti, tra cui costi, velocità e affidabilità.	<i>Lucene</i> : progetto utilizzato per ricerche e text Analytics, incorporato in diversi software open source.

<p><i>Regole associative</i>: tecniche volte a scoprire relazioni interessanti tra variabili all'interno di un grande dataset.</p>	<p><i>Mahout</i>: progetto Apache il cui obiettivo è sviluppare applicazioni gratis per algoritmi di machine learning distribuiti e scalabili, al fine di supportare i Big Data Analytics sulla piattaforma Hadoop.</p>
<p><i>Regressione</i>: set di tecniche che permettono di determinare come il valore di una variabile dipendente cambia quando una o più variabili indipendenti vengono modificate.</p>	<p><i>MapReduce</i>: software framework introdotto da Google e implementato anche in Hadoop per processare grandi Data set relativamente a certi tipi di problemi su un sistema distribuito.</p>
<p><i>Sentiment analysis</i>: applicazione del processing natural language e di altre tecniche analitiche per identificare ed estrarre informazioni soggettive dai testi, per esempio la "polarità" (positiva, negativa o neutra) delle caratteristiche o dei prodotti su cui le persone hanno espresso un'opinione e il grado e la forza dell'opinione stessa.</p>	<p><i>MPP Database</i>: Database che forniscono un'interfaccia SQL e lavorano segmentando su più nodi i dati i quali vengono elaborati in parallelo</p>
<p><i>Statistica</i>: scienza della raccolta, organizzazione e interpretazione dei dati, utilizzata per esprimere giudizi sulle relazioni tra variabili che potrebbero essersi verificate per caso (ipotesi nulla) e su quelle causali (statisticamente significative).</p>	<p><i>Oozie</i>: progetto open source che ottimizza il flusso di lavoro e il coordinamento tra i task.</p>
<p><i>Visualizzazione</i>: tecniche di creazione di immagini, diagrammi o animazioni che consentono di comunicare, capire e migliorare i risultati dell'analisi dei Big Data.</p>	<p><i>PIG</i>: altro software sviluppato da Yahoo e ora open source, che cerca di portare Hadoop più vicino alla realtà di sviluppatori e utenti aziendali, analogamente a Hive. A differenza di quest'ultimo però utilizza un linguaggio "Perl-like" per eseguire query sui dati memorizzati</p>

	in un cluster Hadoop.
	<i>Sistemi distribuiti:</i> insieme di computer che comunicano attraverso una rete, utilizzati per risolvere un problema computazionale, il quale viene diviso in diversi task ognuno dei quali viene risolto da uno o più computer che lavorano in parallelo. Questi sistemi presentano tre vantaggi: costi bassi, alta affidabilità e una maggiore scalabilità.
	<i>WibiData:</i> combinazione di web Analytics e Hadoop, costruito su HBase. Permette ai siti web di eseguire esplorazioni migliori e di lavorare con i dati dei propri utenti, consentendo risposte in real-time relativamente al comportamento degli utenti e fornendo quindi contenuti personalizzati e decisioni.
	<i>Zookeeper:</i> infrastruttura centralizzata con vari servizi, in grado di sincronizzare un cluster di servers. Le applicazioni di Big Data Analytics si servono di questi servizi per coordinare processi paralleli tra grandi cluster.

Tabella 2: Tecniche e tecnologie per l'analisi dei Big Data

Dopo aver descritto le principali tecniche e tecnologie, vediamo ora un'architettura concettuale per l'analisi dei Big Data, rappresentata nella figura riportata nella pagina seguente.

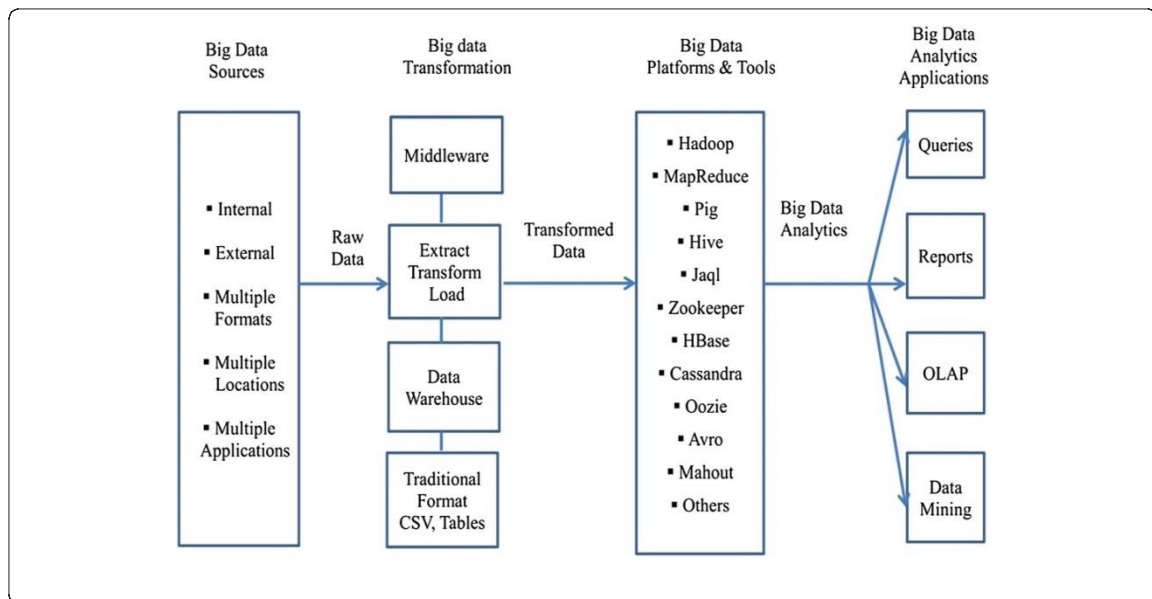


Figura 3: Architettura concettuale per l'analisi dei Big Data (Raghupathi W., 2014)

Innanzitutto i Big Data provengono da diverse fonti, sia interne che esterne, spesso sono in formati differenti e risiedono in posizioni multiple in numerosi sistemi legacy e altre applicazioni. I dati possono essere sia strutturati (dati conservati in Database relazionali, organizzati secondo schemi e tabelle rigide), sia non strutturati (dati conservati senza alcuno schema come forme libere di testo tra cui articoli e parti di e-mail, audio senza tag, immagini e video) sia semi-strutturati (dati che presentano caratteristiche sia di quelli strutturati che di quelli non strutturati; un esempio è rappresentato dai file compilati con sintassi XML per i quali non ci sono limiti strutturali all'inserimento dei dati, ma le informazioni vengono organizzate secondo logiche strutturate e interoperabili). Dopo che i dati sono stati uniti, questi hanno bisogno di essere processati o trasformati, essendo in uno stato grezzo. Ci sono diverse opzioni a disposizione:

- Service-oriented architecture combinata con web services (middleware): i dati rimangono grezzi e i services sono utilizzati per chiamare, recuperare e processare i dati;
- Data warehousing: dati provenienti da svariate fonti vengono aggregati e preparati per essere processati, anche se non sono disponibili in real-time;
- Extract, transform and load (ETL): dati che derivano da diversi fonti vengono puliti e preparati per lo step successivo.

Il passo successivo consiste nella scelta della piattaforma e della tecnologia da utilizzare, tra quelle elencate nella tabella.

L'ultima fase invece è relativa alle applicazioni di Big Data Analytics che includono queries, reports, OLAP e Data mining e alla visualizzazione, compresa in tutte queste applicazioni (Raghupathi W., 2014).

Un ruolo centrale in quest'ambito viene svolto dai Big Data Analytics, tecnologie di Business Intelligence & Analytics basate sulle tecniche descritte in tabella. Riportiamo nel seguente grafico quelli più importanti, che ritroveremo più volte in questa ricerca.



Grafico 2: I Big Data Analytics

- *Behavioural/Gestural Analytics*: analisi automatizzata delle attività umane catturate da video che tracciano i movimenti e i gesti per individuare e comprendere comportamenti e intenzioni;
- *Content Analytics*: insieme di tecnologie che processano i contenuti digitali e i comportamenti degli utenti nelle conversazioni con altre persone, nelle discussioni sui social network o relativamente al livello di consumo ed engagement di documenti e nuovi siti, per rispondere a determinate domande;
- *CRM Analytics*: soluzioni che raccolgono, organizzano e sintetizzano i dati dei clienti per aiutare le organizzazioni a risolvere i problemi di business riguardanti i consumatori

attraverso tool, dashboard, portali e altri metodi negli ambiti di Marketing, Sales e Customer Service;

- *Customer Analytics*: tecnologie che sfruttano i dati per capire la composizione, i bisogni e la soddisfazione dei consumatori, per poi segmentarli in gruppi sulla base dei comportamenti adottati, implementare azioni di Marketing personalizzate e determinare trend generali;
- *Descriptive Analytics*: analisi di dati e contenuti per rispondere alla domanda “Cosa è successo?” o “Cosa sta succedendo?” attraverso strumenti tradizionali di BI e visualizzazione;
- *Predictive Analytics*: Analytics avanzati che implementano tecniche quali la regressione, i modelli predittivi e la statistica per analizzare i dati e i contenuti e rispondere alle domande “Cosa succederà” o “Cosa accadrà molto probabilmente?”;
- *Prescriptive Analytics*: altra forma di Analytics avanzati che esamina i dati e i contenuti per rispondere alle domande “cosa dovrebbe essere fatto?” o “cosa dobbiamo fare per far sì che succeda una determinata cosa?” e per far questo utilizza tecniche quali l’analisi di grafici, la simulazione, le reti neurali e la machine learning;
- *Social Analytics*: tools che estraggono, analizzano e sintetizzano automaticamente i contenuti generati dagli utenti online. Questa tecnologia verrà descritto in modo approfondito nel successivo capitolo;
- *Text Analytics*: processo di estrazione delle informazioni dai testi, utilizzato per diversi scopi, tra cui il *riepilogo*, ovvero il tentativo di trovare i contenuti chiave in un grande insieme di informazioni, la *sentiment analysis*, già spiegate o per determinare cosa ha guidato un determinato commento di una persona e quindi per un fine esplicativo;
- *Web Analytics*: applicazioni analitiche utilizzate per capire e migliorare l’esperienza online del consumatore, l’acquisizione di utenti e l’ottimizzazione del digital Marketing delle campagne pubblicitarie. Questi offrono reporting, segmentazione, gestione delle campagne e integrazione con altre fonti dati e processi.

## 1.5 TENDENZE IN ATTO

La presenza di nuovi strumenti e i miglioramenti degli approcci di analisi dei dati offrono nuove opportunità per trarre ulteriori vantaggi. In particolare possiamo individuare tre tendenze in atto (Court D., 2015).

La prima è rappresentata dallo sviluppo di soluzioni analitiche specifiche, focalizzate su un’area determinata, come la logistica, la gestione del rischio, il pricing e la gestione del personale. Si

tratta di soluzioni che possono essere sviluppate molto rapidamente e che permettono di ottenere notevoli benefici nell'area specifica, ma che richiedono un cambiamento nella cultura organizzativa e la creazione di enfasi sulla loro adozione.

Un second trend è la democratizzazione degli "Analytics", su cui stanno investendo molte aziende quali American Express, Procter & Gamble e Walmart (Court D., 2015). Lo sviluppo di nuovi strumenti self-service sta aumentando la fiducia negli Analytics degli utilizzatori frontline, i quali, senza conoscere la singola riga di codice, possono collegare i dati da molteplici fonti e fare previsioni. Inoltre gli strumenti di visualizzazione rendono più facili le operazioni di slice e dice, permettono di individuare i dati da esplorare per affrontare le problematiche di business e di supportare il processo di decision making. Un'azienda di hardware, per esempio, ha sviluppato un set di soluzioni analitiche self-service e di strumenti di visualizzazione che aiutano l'azienda a condurre le analisi dei clienti e a identificare le opportunità di vendita e di rinnovo al fine di migliorare le decisioni della forza di vendita. L'implementazione di questa piattaforma ha portato ad un incremento dei ricavi pari a 100 milioni di dollari.

Infine oggi sta diventando sempre più semplice automatizzare i processi e prendere decisioni: i miglioramenti tecnologici permettono di catturare un numero molto più grande di dati in tempo reale, facilitando quindi i processi di elaborazione di un'enorme base di dati e di analisi in real time. Questi avanzamenti tecnologici stanno aprendo nuovi percorsi per l'automazione e l'apprendimento automatico per tutte le aziende e non solo quelle leader nella tecnologia. Per esempio un'importante società di assicurazione ha raggiunto notevoli progressi implementando uno strumento che permette di prevedere la gravità dei reclami, attraverso un confronto istantaneo di milioni di dati registrati, riducendo quindi il bisogno dell'intervento umano.

## **1.6 L'IMPATTO DEI BIG DATA IN QUATTRO SETTORI**

Affrontiamo ora il modo in cui i Big Data possono creare valore in quattro diversi domini: il *settore sanitario*, il *settore della Pubblica Amministrazione*, il *Manufacturing* e il *Retail*.

### **1.6.1 SANITÀ**

Oggi il settore sanitario sta affrontando uno tsunami di dati relativi alla salute e alla Sanità, creati e accumulati continuamente. Si tratta di dati clinici generati da sistemi di supporto alle decisioni quali note e prescrizioni dei medici, immagini mediche comprese quelle 3D più recenti, dati dai laboratori, dalle farmacie, dalle assicurazioni e altri amministrativi; electronic

health records (EHR, archivi di dati digitali sanitari); dati generati dalle macchine e sensor data, come quelli provenienti dal monitoraggio dei segnali vitali; genomic data, tra cui il genotipo e l'espressione genica; post dai social media; blog (Raghupathi W., 2014). In particolare sono gli EHR che consentono una profonda conoscenza clinica e dei quadri patologici dei pazienti: questi archivi di dati possono essere utilizzati per cercare associazioni nelle diagnosi mediche e considerare le relazioni temporali tra eventi al fine di scoprire la progressione delle malattie (Chen H., 2012). Ogni persona porta 4 terabyte di dati e ciascun payer-provider (i payors sono compagnie di assicurazione, organizzazioni sanitarie, imprenditori e gestori di richieste di rimborso nell'ambito dei programmi di assistenza medica statali o federali mentre i providers sono ospedali, personale sanitario e cliniche) potrebbe costruire una matrice con centinaia di migliaia di pazienti con diverse informazioni e parametri (demografia, cura e risultati) raccolti per un lungo periodo di tempo (Miller K., 2011-2012).

I dati sono accumulati velocemente in tempo reale. Future applicazioni real-time, come il rilevamento di infezioni il prima possibile, permetteranno di identificarle rapidamente e di mettere in pratica le giuste terapie, prevenendo infezioni e riducendo il tasso di mortalità.

Anche la varietà è una caratteristica dei dati del settore. Ci sono infatti dati strutturati che possono essere facilmente memorizzati, interrogati, analizzati e manipolati, i quali, insieme a quelli semistrutturati, includono la lettura degli strumenti e la conversione dei documenti cartacei in EHR. Oltre a questi ci sono dati non strutturati che comprendono note di medici e infermieri scritte a mano, prescrizioni cartacee, radiografie e immagine mediche.

La veridicità dei dati, soggetti a errori (soprattutto quelli non strutturati caratterizzati da grande variabilità), rappresenta un obiettivo, in quanto la qualità dei dati nella Sanità è fondamentale: decisioni mediche che impattano sulla vita o la morte delle persone dipendono infatti dall'accuratezza delle informazioni (Raghupathi W., 2014).

Le principali applicazioni di Big Data Analytics utilizzate nel settore sanitario sono le queries, i report, l'OLAP e il Data mining. Tecniche e tecnologie, di cui abbiamo già discusso in precedenza, vengono utilizzate per aggregare, manipolare, analizzare e visualizzare tutti i dati.

Per riuscire ad ottenere vantaggio competitivo dagli insight forniti dai Big Data, gli strumenti tradizionali non sono sufficienti in quanto essi si focalizzano solo sulla riduzione dei costi e non sul miglioramento dei risultati delle terapie e quindi delle condizioni e del livello di soddisfazione dei pazienti. È quindi necessario sviluppare strumenti incentrati sul paziente che



prendano in considerazione entrambi gli aspetti; per questo gli stakeholder devono focalizzarsi sui cinque seguenti aspetti:

- right living: i pazienti devono avere un ruolo attivo nel miglioramento della loro salute, facendo scelte adeguate relative alla dieta, all'esercizio e alla prevenzione;
- right care: i medici e tutto il personale sanitario devono avere accesso alle medesime informazioni per favorire il coordinamento e lavorare per lo stesso obiettivo, al fine di evitare la duplicazione dello sforzo e strategie subottimali;
- right provider: tutti i provider devono accedere agli archivi di dati ed essere in grado di raggiungere i risultati migliori;
- right value: incrementare contemporaneamente il valore e la qualità della cura;
- right innovation: sviluppare nuovi approcci per migliorare i servizi sanitari.

Già alcuni leader nel settore sanitario hanno iniziato a focalizzarsi su queste soluzioni o comunque a porre le basi per il futuro. Per esempio Kaiser Permanente, consorzio sanitario in Oakland, California, United States, ha implementato HealthConnect, un nuovo sistema che permette lo scambio dei dati tra le varie strutture mediche e promuove l'utilizzo degli EHR; questa soluzione ha migliorato le prestazioni nell'ambito delle malattie cardiovascolari e comportato un risparmio pari a 1 miliardo di dollari dovuto alla riduzione delle visite e dei test di laboratorio (Kayyali B., 2013).

#### 1.6.1.1 Benefici conseguibili dall'utilizzo dei Big Data

Il ricorso ai Big Data in questo settore permette di conseguire una serie di benefici, connessi a quelli dei Big Data in generale. In tabella riportiamo i 6 principali (Manyika J., 2011), (Raghupathi W., 2014), (Sun J., 2013).

Benefici	Descrizione
Riduzione dei costi	L'aumento dell'efficienza è uno dei più importanti vantaggi ed è reso possibile da una serie di pratiche. Innanzitutto la comparative effectiveness research (CER), ovvero l'analisi di grandi Dataset in cui sono contenuti dati quali le caratteristiche dei pazienti e i costi e i risultati delle terapie, permette di identificare trattamenti economicamente

	<p>vantaggiosi. In secondo luogo la creazione di mappe dei processi e di dashboard a partire dai Dataset dei provider consente di identificare le fonti di variabilità e gli sprechi e quindi di ottimizzare i processi clinici. Infine l'analisi dei quadri patologici e dei trend per stimare la domanda futura dei farmaci e il ricorso a modelli predittivi sviluppati dalle aziende farmaceutiche tramite l'aggregazione dei dati di ricerca, favoriscono un'allocazione più vantaggiosa delle risorse R&amp;D.</p>
<p>Incremento dei risultati grazie a decisioni migliori (maggiore qualità delle terapie ed alta soddisfazione dei clienti)</p>	<p>La CER e la trasparenza dei dati, garantita dallo sviluppo di mappe di processo e di dashboard, permettono inoltre di ridurre l'incidenza delle terapie dannose e di quelle che dovrebbero essere prescritte ma che non vengono messe in pratica, nonché di migliorare la qualità delle cure. Anche i sistemi di supporto per le decisioni cliniche, in cui i medici inseriscono le loro terapie che vengono poi confrontate con le linee guida, mettono in guardia da eventuali errori come la prescrizione di farmaci che avranno effetti negativi sui pazienti. Questi ultimi possono essere individuati anche attraverso l'analisi dei test clinici e quindi si eviterà di immetterli sul mercato salvaguardando l'immagine aziendale. Gli stessi modelli predittivi rendono le attività di R&amp;D sui farmaci più veloci e più specifiche, favorendo un'immissione molto più rapida dei medicinali sul mercato e la produzione di composti specifici con alto tasso di successo. Infine l'utilizzo di un Database con i dati di tutti i pazienti e di tutte le terapie a livello nazionale assicura una rilevazione rapida delle malattie infettive e il controllo di eventuali epidemie globali grazie ad un apposito programma, garantendo quindi il monitoraggio della salute pubblica ed un miglioramento della qualità della vita stessa.</p>

Personalizzazione della cura	Enormi Dataset vengono analizzati per esaminare le relazioni tra le variazioni genetiche, la predisposizione a malattie specifiche e le risposte a determinati farmaci per poter sviluppare farmaci personalizzati, garantendo terapie più efficaci e diagnosi precoci.
Prevenzione	L'utilizzo di Analytics avanzati quali modelli predittivi e di segmentazione dei pazienti per identificare gli individui che possono trarre beneficio da cure proattive o da cambiamenti dello stile di vita, favorisce la prevenzione di eventuali problemi di salute.
Aumento del livello di soddisfazione di tutti gli attori del sistema	Le piattaforme e le comunità online, come i già citati Sermo.com e PatientsLikeMe.com, migliorano la comunicazione tra gli individui e consentono la condivisione di esperienze e quindi un aumento del sostegno reciproco e del loro coinvolgimento.
Sviluppo di nuovi modelli di business	I dati clinici dei pazienti e dei Dataset delle richieste, aggregati e sintetizzati, possono essere venduti a terze parti. Questi robusti Dataset clinici rendono possibile la nascita di nuovi business, quali l'analisi dei risultati clinici per i payors che possono così prendere decisioni migliori o l'analisi dei Database per scoprire i biomarcatori che selezionano le terapie.

*Tabella 3: Benefici conseguibili dall'utilizzo dei Big Data nel settore sanitario*

È quindi chiaro che l'estrazione di conoscenza dai Big Data nel settore sanitario sia fonte di rilevanti vantaggi, tuttavia essa richiede di affrontare sfide quali la garanzia della privacy in modo da mantenere la fiducia delle persone e la conduzione di ricerche etiche. Il rischio è infatti quello di essere puniti per il mancato rispetto dei requisiti stabiliti dall'HIPAA (Health Insurance Portability and Accountability Act) e dall'IRB (Institutional Review Board).

## 1.6.2 PUBBLICA AMMINISTRAZIONE

I Governi di tutto il mondo stanno riconoscendo l'importanza dei Big Data come leva per migliorare i servizi ai cittadini e ottenere un vantaggio competitivo, soprattutto in un periodo come quello attuale, in cui i Governi locali stanno affrontando un calo del budget, un deterioramento delle infrastrutture e il bisogno di attirare e trattenere business per la crescita economica. Per la Pubblica Amministrazione è fondamentale capire come i Big Data e gli Analytics possano incrementare l'efficienza e la produttività, migliorare i processi decisionali, aumentare la trasparenza verso i cittadini, aiutare a controllare sprechi, frodi e abusi e gestire budget e costi (Fiorenza P., 2014). Cisco nel 2014 ha stimato che lo sfruttamento dei Big Data rappresenterà un'opportunità pari a 4600 miliardi di dollari nei prossimi 10 anni (Bikar P., 2015).

Gli enti organizzativi hanno accesso a vari flussi di dati, per lo più testuali e numerici e per il 90% in formato digitale (Manyika J., 2011): documenti, video, foto, dati dai social media e dal mobile. Si tratta quindi di un mix di dati strutturati e non, che necessitano, per l'estrazione di insight, di un Database NoSQL, più agile e flessibile nel raccogliere, processare e analizzare un'ampia varietà di fonti di dati. MarkLogic è la piattaforma NoSQL collaudata per le applicazioni Big Data governative, realizzate per condurre analisi real-time e trasformare tutti i dati in informazioni di valore e azioni da implementare (Fiorenza P., 2014).

Comunque, nonostante ci siano programmi di e-government avanzati, le agenzie non condividono i dati con i cittadini e con le organizzazioni: molto spesso succede che l'impiegato di un'agenzia riceva una copia di dati via fax o via mail da un'altra, quando questi potrebbero essere memorizzati elettronicamente. Esistono addirittura delle restrizioni legali e delle policy che impediscono la condivisione.

### 1.6.2.1 Benefici conseguibili dall'utilizzo dei Big Data

Analizziamo ora più in dettaglio i benefici apportati dai Big Data alla Pubblica Amministrazione.

La virtualizzazione degli eterogenei Dataset pubblici permette di guidare il processo decisionale e di migliorare i risultati delle politiche. Ci sono però degli ostacoli da superare per perseguire questo beneficio:

- ci sono troppi dati da utilizzare efficacemente;
- i dati si sviluppano ampiamente e in modo nascosto;
- i formati non sono facilmente condivisibili;

- le politiche di sicurezza variano dal dominio pubblico a quello privato e segreto.

La soluzione è rappresentata dal Data Virtualization (DV) che fornisce un'unica soluzione consentendo l'accesso a tutti i dati, senza che ci siano duplicazioni. Per esempio il Governo Tedesco ha implementato un programma decennale che prevede la digitalizzazione di regole e regolamenti riguardanti la pianificazione, le zone edificabili e i permessi ambientali. Il programma si pone come obiettivi decisioni migliori e più veloci, un flusso di lavoro completamente digitale, una visione unica delle informazioni per cittadini, governi e aziende e una condivisione dei dati accurata e affidabile (Bikar P., 2015).

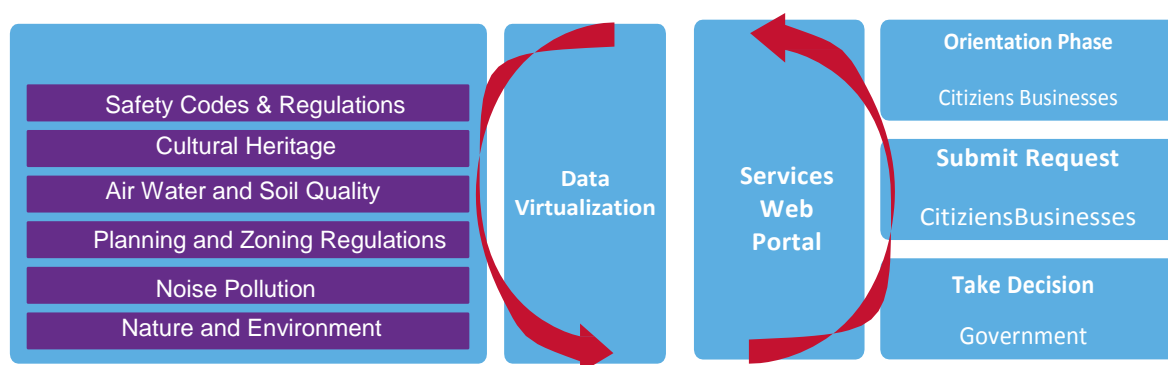


Figura 4: Programma di digitalizzazione di regole e regolamenti implementato dal Governo Tedesco (Bikar P., 2015)

I Big Data danno vita a nuove politiche basate su informazioni più smart relative ai cittadini e alle aziende, rendendo i servizi pubblici più efficaci in diversi modi: garantendo alle famiglie la distribuzione di benefit e di altri supporti di loro diritto, rispondendo alle richieste pubbliche più velocemente e accuratamente, rilevando in anticipo problemi e priorità in modo da aggiustare le politiche. Inoltre prendendo decisioni migliori su come gli enti devono essere organizzati e quali lavori hanno la priorità, i costi delle operazioni governative si riducono.

Rendere le informazioni del Settore Pubblico disponibili a tutti i cittadini e a tutti i dipendenti della PA genera trasparenza, un aumento della soddisfazione e della fiducia dei cittadini stessi nei confronti della PA, facilita l'interoperabilità all'interno dell'amministrazione e la creazione di nuovi servizi.

Analizzando i grandi volumi di dati in real-time e trasformandoli in informazioni di valore tramite i modelli predittivi, è possibile scoprire anomalie passate e attuali, rilevare trend che evidenziano rischi potenziali e trovare quindi dei modi per ridurre errori, abusi e frodi e ottimizzare i processi (Bikar P., 2015).

I Big Data permettono anche di scoprire la variabilità delle performance delle diverse parti di un ente governativo che svolgono funzioni simili, attraverso per esempio l'utilizzo di dashboard che visualizzano tutti i dati operativi e finanziari, alimentando la competizione tra queste; competizione che porta quindi ad un miglioramento delle performance e ad un aumento della produttività.

L'utilizzo dei Big Data Analytics dà anche la possibilità segmentare e personalizzare i servizi per gli individui e la popolazione. I vantaggi sono un incremento dell'efficacia e dell'efficienza delle agenzie, un miglioramento della relazione tra manager e clienti degli enti e un aumento della soddisfazione dei clienti grazie a servizi che rispondono ai loro bisogni. Un'agenzia di lavoro tedesca per esempio ha analizzato i dati storici dei clienti, le azioni implementate e i risultati al fine di effettuare interventi su misura per i disoccupati: l'iniziativa ha portato a una riduzione annua della spesa di 10 miliardi nei tre anni successivi, ad una diminuzione del tempo necessario per trovare lavoro da parte disoccupati e a una loro maggiore soddisfazione (Manyika J., 2011).

### 1.6.3 MANUFACTURING

Anche il Manufacturing sta cercando di applicare Analytics a grandi pool di dati al fine di dedurre informazioni utili per il business e azioni da perseguire. I benefici che ne derivano sono un incremento dell'efficienza nelle attività di progettazione e di produzione, un aumento della qualità e del grado di innovazione dei prodotti, una maggiore soddisfazione dei bisogni dei consumatori, una previsione della domanda più accurata, una produttività più alta e un miglioramento della gestione della complessa e globale catena del valore.

I dati provengono da innumerevoli fonti: macchinari per la produzione, sistemi per la gestione della supply chain, sistemi che monitorano le performance dei prodotti che sono già stati venduti, RFID (radio-frequency identification) ovvero dispositivi che tracciano il prodotto e per i quali è previsto un aumento dai 12 milioni del 2011 ai 209 miliardi del 2021, commenti sui social media e diversi sistemi tra cui i computer-aided design, i computer-aided engineering e i computer-aided manufacturing (Manyika J., 2011).

Segnali quali le vibrazioni e la pressione estratti da sensori presenti nelle macchine e dati storici relativi a questi elementi possono essere utilizzati da Analytics avanzati che con l'avvento del cloud computing e del framework Cyber-Physical Systems, comporteranno la creazione di un sistema informativo della flotta macchine che consentirà a queste ultime di essere self-aware e di prevedere eventuali problemi di prestazione. Un sistema di macchine self-aware e self-

manteined è un sistema in grado di autovalutare il suo stato di salute e di peggioramento ed utilizza ulteriori informazioni provenienti dalle altre macchine per prendere decisioni relative alla manutenzione intelligente al fine di evitare potenziali problemi. In particolare questi smart Analytics verranno utilizzati sia a livello di singola macchina che a livello di flotta (Lee J., 2014).

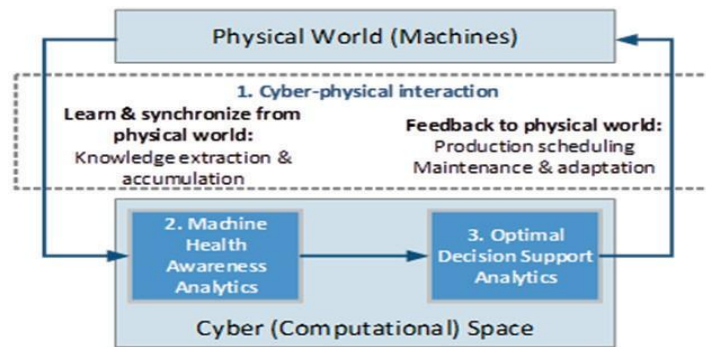


Figura 5: Framework Cyber-Physical Systems per le macchine self-aware e self-maintenance (Lee J., 2014)

Per fare in modo che i team di Big Data e di Analytics riescano ad estrarre la business intelligence dalle diverse fonti dati e ad ottimizzare i processi manifatturieri, devono valere le tre seguenti condizioni (Kerschberg B., 2014):

- *Trasparenza dei dati* che permette l'integrazione dei dati provenienti dalle diverse funzioni manifatturiere;
- *Visibilità del processo* che consente ai manager di vedere come stanno avvenendo i processi, in modo da implementare delle correzioni;
- *Visualizzazione dei dati* risultanti dall'applicazione degli Analytics ai Big Data in modo che gli utilizzatori finali come i responsabili dell'impianto possano vedere i dati nascosti e il loro valore. Questo è fondamentale per il dinamismo real-time dei dati stessi.

Vediamo di seguito i benefici derivanti dallo sfruttamento dei Big Data nelle diverse parti della catena del valore manifatturiera.



- Sfruttando i dati di input e gli insight sui clienti, realizzare un prodotto con caratteristiche adeguate alle loro esigenze e ridurre i costi di sviluppo del prodotto;
- creare più valore attraverso la piattaforma di product lifecycle management (PLM) che integra i Dataset di sistemi multipli per rendere efficace la collaborazione;
- prendere le decisioni migliori, che comportano una riduzione dei costi: produttori e designer condividono i dati e creano in modo rapido ed economico simulazioni per testare diversi progetti, diversi fornitori e i costi di produzione associati (importante perché le decisioni nella parte di progettazione comportano l'80% dei costi);
- ridurre il tempo di sviluppo di nuovi prodotti ed eliminare i difetti prima della costruzione del prototipo;
- migliorare i prodotti esistenti e sviluppare nuovi modelli e varianti di prodotti già esistenti;
- aumentare il grado di innovazione: i produttori invitano gli stakeholder esterni a esprimere idee innovative o a partecipare allo sviluppo di nuovi prodotti attraverso piattaforme web-based.

- Prevedere la domanda e migliorare la pianificazione della supply chain, in modo da utilizzare i soldi nel modo più efficiente possibile e migliorare il livello di servizio. Per far questo è necessario integrare i dati dei Retailer come quelli relativi alle promozioni (item, prezzi,..) e al magazzino (livello di scorte nel magazzino, vendite per negozio);
- ridurre il tempo di risposta, il livello delle scorte e il tempo per lanciare un nuovo prodotto (da articolo introduttivo);
- estrarre nuove idee e comprendere meglio i propri prodotti, clienti e mercati (da articolo introduttivo).

- Aumentare l'efficienza del processo produttivo attraverso tecniche di simulazione applicate ai grandi volumi di dati generati dai prodotti;
- ridurre il numero di cambiamenti del progetto, i costi dei tool di design e costruzione, le ore di assemblaggio e migliorare l'affidabilità della consegna. Per ottenere tali vantaggi è necessario utilizzare i dati di sviluppo del prodotto e i dati storici di produzione per progettare e simulare il sistema produttivo, dal layout alla sequenza di fasi da seguire per un dato prodotto;
- controllare e ottimizzare i processi produttivi e di supply al fine di ridurre gli scarti e



<p>massimizzare la resa utilizzando i dati real-time dei sensori in tali processi.</p>
<ul style="list-style-type: none"> <li>• Riprogettare il prodotto e sviluppare nuovi prodotti, sfruttando i dati appena citati;</li> <li>• migliorare la previsione della domanda;</li> <li>• migliorare i servizi post-vendita offerti (esempio: i produttori di aerei o di ascensori utilizzano i dati dai sensori presenti sui prodotti per programmare pacchetti proattivi di servizi di manutenzione intelligente).</li> </ul>

Tabella 4: Benefici derivanti dallo sfruttamento dei Big Data nelle diverse parti della catena del valore manifatturiera

#### 1.6.4 RETAIL

I Big Data offrono enormi opportunità anche al mondo Retail che ha disposizione non solo dati come le transazioni e le operazioni dei clienti, ma anche dati dagli RFID e informazioni sul comportamento online e sul sentiment dei consumatori. Le aziende che utilizzano tecniche e tecnologie in grado di sfruttare questi dati, riescono a migliorare l'efficacia delle loro azioni di Marketing e di merchandising, a ridurre i costi delle operazioni e della supply chain e quindi a migliorare la loro profittabilità e a ottenere un vantaggio competitivo rispetto agli altri concorrenti (Manyika J., 2011).

I Retailer in particolare possono utilizzare due diverse tipologie di Analytics (Osservatorio Big Data Analytics & Business Intelligence, maggio 2015):

- *Performance Management & Basic Analytics*: strumenti di Descriptive Analytics che consentono di accedere ai dati secondo viste logiche flessibili e dinamiche e di visualizzare sinteticamente e graficamente i principali indicatori di prestazione;
- *Advanced Analytics*: strumenti avanzati che permettono di svolgere un'analisi attiva dei dati sfruttando metodologie di prescriptive&predictive analysis, determinando trend e prevedendo il valore futuro di variabili numeriche e categoriche.

In quest'analisi ci concentriamo sia sui processi di back-end, ovvero i processi di interazione Retail-fornitore o processi interni del Retailer sia su quelli di front-end, ovvero quelli di interazione Retailer-consumatore.

##### ➤ Processi di back-end

- *Logistica.* Oggi i fornitori logistici gestiscono un enorme flusso di beni creando nel contempo un set rilevante di dati (origine e destinazione dei viaggi, dimensione, peso, contenuto del trasporto, posizione) considerati i milioni di viaggi intrapresi ogni giorno.

Le tecniche di Big Data possono essere innanzitutto utilizzate per ottimizzare i costi del viaggio per la consegna dei prodotti. Focalizziamoci su due diverse modalità.

1) *Ottimizzazione del percorso in real time.* Quando il veicolo viene caricato per partire, il percorso di consegna ottimale viene pianificato utilizzando i dati di spedizione rilevati dai sensori sui prodotti in viaggio. Durante il trasporto sistemi dinamici suggeriscono cambiamenti del percorso a seconda delle condizioni del traffico, dei fattori geografici e dello stato del ricevente. Questo approccio, che si basa sull'utilizzo di dati reali, permette di tagliare i costi e di ridurre l'emissione di CO<sub>2</sub> diminuendo per esempio le distanze da percorrere.

2) *Pick-up e consegna basate sulle persone.* Pendolari, taxisti o studenti potrebbero essere pagati per occuparsi della consegna nell'ultima parte del percorso se questa coincide con il viaggio che devono effettuare. Questo approccio, che porta ad una notevole riduzione dei costi, richiede l'utilizzo di tecniche di Big Data: flussi di dati real-time vengono tracciati al fine di assegnare la spedizione alle persone disponibili, basandosi sulle loro posizioni e destinazioni. Attraverso dispositivi mobile, possibili trasportatori pubblicano la loro posizione e accettano l'assegnazione della consegna.

Analytics avanzati possono essere utilizzati anche per prevedere la domanda al fine ottimizzare la capacità di trasporto e la quantità di personale necessario in ciascuna zona. Per l'allocazione delle risorse vengono utilizzate informazioni real-time sulle spedizioni (item che sono appena entrati nella rete distributiva, che sono in transito o che sono in magazzino), informazioni dai clienti (apertura di nuove industrie, fallimenti inaspettati) e anche informazioni di eventi locali (epidemie regionali e disastri naturali).

Le soluzioni Big Data permettono inoltre il recupero di informazioni utili per il rilevamento dei rischi di supply chain: dati sugli sviluppi politici locali, sull'economia, sulla salute, sulla natura provenienti da diverse fonti quali siti, social media, blog vengono aggregati e analizzati attraverso semantic Analytics e

altre tecniche. Queste soluzioni individuano dei pattern tra le diverse informazioni e quando si verifica una condizione critica per la supply chain del cliente, questo viene avvisato e gli viene inviato un report contenente informazioni quali la probabilità e l'impatto del rischio e contromisure per mitigarlo (Jeske M., 2013).

- *Gestione del magazzino.* L'integrazione di strumenti di ottimizzazione del magazzino e dei sistemi ERP consente di fare la migliore analisi possibile dei dati a disposizione quali dati delle vendite, degli acquisti, finanziari, di fornitura e di produzione. Sviluppando un algoritmo, il sistema di ottimizzazione può creare un'interfaccia grafica in grado di illustrare una sintesi di tutti i dati, che permette di identificare i cambiamenti stagionali dei prodotti richiesti, quando si verificheranno gli stock-out, le vendite perse e gli ordini in eccesso, ecc... Questa analisi viene quindi utilizzata da uno tool di ottimizzazione del magazzino per prevedere la domanda in modo accurato e stabilire quindi il livello ottimale del magazzino, che soddisfi le richieste dei clienti ed eviti sia lo stock-out che la sovrabbondanza di item. Questo strumento dà poi dei suggerimenti per la successiva pianificazione delle scorte e definisce la soglia del riordino. Data l'integrazione tra questo tool e i sistemi ERP, tutti questi dati vengono comunicati sia all'interno che agli stakeholder all'esterno dell'organizzazione (Sage, 2013) Alcune aziende utilizzano anche sistemi bar-code collegati ai processi di rifornimento automatico per ridurre l'incidenza dello stock-out (Manyika J., 2011).

- *Pianificazione dei turni.* È possibile utilizzare un algoritmo predittivo che prenda in considerazione un ampio range di parametri individuali e locali: dati dei ricavi storici, orari di apertura dei negozi, orari di arrivo dei prodotti dai centri distributivi ma anche i giorni del mercato, i giorni di vacanza delle località vicine e i dati delle previsioni meteorologiche che influenzano il comportamento dei clienti. L'algoritmo dà come soluzione le vendite giornaliere previste, a partire dalle quali vengono pianificati i turni in modo ottimale, evitando surplus e carenze del personale, che impattano negativamente sulle performance finanziarie del negozio e sulla soddisfazione dei clienti e dei dipendenti (Jeske M., 2013).

Le nuove abitudini dei consumatori, che comportano la creazione di uno tsunami di dati, giustificano l'utilizzo delle tecnologie di Big Data da parte dei Retailer (Osservatorio Big Data Analytics & Business Intelligence, maggio 2015):

- 2 consumatori su 3 si informano online prima di acquistare un prodotto, comprano in un negozio, ma hanno preso la decisione prima su canali digitali (Fonte: Net Retail –Netcomm, Campione: 3.055 individui);
- il 54% dei consumatori preferisce percorsi di acquisto che contemplino almeno un'interazione con i canali digitali online e mobile (Fonte: Total Retail PwCCampione:1.000 consumatori che hanno acquistato almeno 1 volta online);
- degli oltre 30 milioni di Internet user, il 31% utilizza per navigare un solo device, il 69% utilizza due o più device per navigare tra cui PC, Smartphone e Tablet (Fonte: SurveyCAPI 2013 –Doxa, Campione: totale Internet userdaily);
- l'89% degli utenti Smartphone utilizza il device all'interno del negozio. Di questi circa il 40% dichiara di farlo sempre o spesso (Fonte: SurveyCAWI 2014 –Doxa, Campione: 1.503 utenti smartphone);
- tra gli utenti che usano lo Smartphone in negozio, il 42% degli utenti confronta i prezzi e il 30% degli utenti invia foto dei prodotti da acquistare ad amici (Fonte: SurveyCAWI 2014 –Doxa, Campione: 1.341 utenti che usano lo smartphonein punto vendita);
- molti consumatori utilizzano applicazioni quali RedLaser che permette loro di scannerizzare il bar code su un item in un negozio con il loro smartphone e ottenere immediatamente il prezzo e il prodotto dei concorrenti (Manyika J., 2011).

I Retailer hanno a disposizione una serie di leve per sfruttare i Big Data e trarre quindi vantaggio nei processi di interazione con il cliente. In questo paragrafo ci limitiamo ad elencarle, in quanto saranno oggetto di approfondimento del secondo capitolo. Nei processi di Marketing le tra le leve principali citiamo il Cross-Selling, il location based Marketing, l'in-store behavior analysis, la customer micro-segmentation, la sentiment analysis e la multichannel consumer experience; mentre nelle attività di merchandising l'ottimizzazione dell'assortimento, l'ottimizzazione del prezzo e l'ottimizzazione del posizionamento e del design (Manyika J., 2011). L'utilizzo di queste tecniche impatta positivamente sulla customer experience del consumatore, dalla fase di prevendita a quelle di acquisto e di pagamento fino alla fase di post-vendita: il cliente riesce a

trovare il prodotto soddisfa al meglio i suoi bisogni e spende meno per trovare i prodotti al prezzo più vantaggioso. Anche su questo argomento ci focalizzeremo nel capitolo successivo.

## 1.7 GOVERNANCE DEI SISTEMI BDA&BI

Affinché le aziende riescano a sfruttare i Big Data, è necessario che al loro interno ci siano delle figure che si occupino della governance dei sistemi di Big DataAnalytics& Business Intelligence (BDA&BI). I due ruoli principali sono quelli del *Data Scientist* e dello *Chief Data Officer* (CDO), che descriviamo nel dettaglio di seguito.

### 1.7.1 DATA SCIENTIST

Il Data Scientist rappresenta il lavoro più affascinante (Ariker M. M. T., 2013) del ventunesimo secolo e, come abbiamo già visto in precedenza, la domanda di questa figura professionale è molto alta.

Ma chi è il Data Scientist? Il Data Scientist è una persona con una solida formazione in computer science, modellazione, statistica, matematica e Analytics, dotata di un forte senso nel business e una grande abilità nel comunicare le sue scoperte ai leader dell'azienda tanto da influenzare l'approccio dell'organizzazione nell'affrontare le sfide di business. Egli esplora i dati da molteplici fonti al fine di ricavare insight nascosti che, trasformati, possono fornire un vantaggio competitivo o risolvere un problema urgente. Non si limita a cercare e riferire i dati, ma li osserva da diverse angolature, determinando il loro significato e scovando dei modi per applicarli. Per questo deve avere una conoscenza, almeno di base, della strategia aziendale (IBM).

In particolare, il Data Scientist deve sviluppare le seguenti 8 competenze (Holtz D., 2014):

- *basic tools*: saper utilizzare un linguaggio di programmazione statistica come R e un linguaggio di querying del Database come SQL;
- *basics statistic*: essere familiari con testi statistici, distribuzioni, stimatori di probabilità e capire quali tecniche rappresentano un approccio valido;
- *machine learning*: conoscere le diverse tecniche di machine learning e capire quando implementare l'una o l'altra;
- *Data munging*: sapere come affrontare le imperfezioni dei dati, quali valori mancanti, formattazione errata di stringhe o dati;

- *Data visualization & communication*: essere in grado di descrivere gli insight ad altre persone sia in modo tecnico che non ed avere una certa pratica con gli strumenti di visualizzazione e dei principi alla base della codifica dei dati;
- *calcolo multivariato e algebra lineare*: essere in grado di risolvere calcoli multivariati o problemi di algebra lineare, dato che rappresentano la base delle tecniche di machine learning;
- *software engineering*: avere un forte background di software engineering ed essere responsabili della gestione della registrazione dei dati e dello sviluppo dei prodotti guidato dai dati;
- *pensare come un Data scientist*: sentirsi delle persone in grado di risolvere i problemi, basandosi sui dati.

Oltre a queste competenze che possono essere apprese in Università, il Data Scientist deve possedere anche specifici tratti di personalità: deve essere curioso, desideroso di analizzare in profondità i dati per trovare delle risposte a problemi ancora sconosciuti, sicuro di sé stesso e paziente nel ricavare insight da una massa di dati, che potrebbe richiedere molto tempo e molti tentativi (Van Rijmenam M., Data Scientist Profile, 2015) .

### 1.7.2 CHIEF DATA OFFICER

Riconosciuto l'impatto che i dati hanno sulle performance e sull'innovazione, i dirigenti delle aziende hanno compreso l'importanza di avere una persona al loro interno che svolga il ruolo di leader dei dati e guidi una crescita basata sui dati stessi. Questa persona è lo Chief Data Officer (CDO). Il CDO è il responsabile della strategia dei Big Data all'interno dell'azienda: egli definisce, sviluppa e implementa i metodi attraverso i quali l'organizzazione acquisisce, gestisce e governa i dati e cerca di identificare nuove opportunità di business sfruttando i dati in modo creativo. Data la strategicità assunta dai dati, il CDO dovrebbe essere un membro del Consiglio e rispondere direttamente al CEO.

Lo Chief Data Officer ha a disposizione cinque diversi modi per guidare l'innovazione e la crescita (Teerlink M., 2014):

- *Data leverage*: prevede di ricavare informazioni rilevanti e insight di valore dai dati esistenti fine di aumentare l'efficienza e la produttività, migliorare l'immagine, incrementare i ricavi o innovare per differenziarsi dai competitor;

- *Data enrichment*: consiste nell'ampliamento dei Dataset attraverso la combinazione di fonti di dati interne, l'acquisizione di dati esterni dai governi o dai social media e l'integrazione di dati dei business partners.
- *Data monetization*: si focalizza sulla scoperta di nuove opportunità di guadagno e ricavo, che vanno ad impattare sul modello di business esistente e sulla strategia organizzativa.
- *Data upkeep*: si riferisce alla gestione e alla garanzia della qualità dei dati, considerata la loro crescente eterogeneità.
- *Data protection*: è un elemento molto importante visto l'alto rischio associato al fallimento di proteggere i dati come un asset. Per proteggere i dati, il CDO collabora con il responsabile della sicurezza delle informazioni, lo Chief Information Security Officer (CISO).

Il CDO deve inoltre garantire trasparenza relativamente ai dati utilizzati e la privacy: i clienti devono essere consapevoli dei dati che sono stati raccolti su di loro e avere la possibilità di eliminarli nel caso in cui non siano stati memorizzati in modo anonimo.

Il leader dei dati deve inoltre assicurarsi che tutti i dati siano disponibili e vengano condivisi tra le varie unità organizzative, assicurando quindi una maggiore efficienza e innovazione.

Lo Chief Data Officer deve infine colmare il gap tra IT e business: essendo i Big Data una leva per il Marketing e la strategia, con il supporto dall'IT, il CDO deve essere in grado di fare gli interessi di tutti (Van Rijmenam M., Chief Data Officer Profile, 2015).

Il CDO deve possedere un giusto mix di competenze tecniche, di business e sulle persone per guidare il processo di cambiamento e la collaborazione in tutta l'organizzazione, a differenza del Data Scientist che invece ha un background come matematico e statistico. Uno Chief Data Officer di successo non deve solamente capire il settore e il mercato in cui opera l'azienda, ma deve anche avere una conoscenza tecnica dei dati, della loro struttura e del loro potenziale come asset. In particolare deve fornire conoscenze di business necessarie per il conseguimento della vision e della strategia, identificare le opportunità generate dai dati e ha delle responsabilità sulla gestione del bilancio per tutte le iniziative relative ai Big Data. Deve anche avere competenze di leadership e di negoziazione per essere in grado di guidare e gestire team di Data scientist e mantenere una buona collaborazione con le altre parti dell'organizzazione che si occupano dei dati. Infine il CDO deve essere dotato di skill tecniche per controllare il flusso dei dati, le fonti dei dati, le capacità dei Data vendor e altre questioni riguardanti i dati

stessi e avere una certa esperienza con le soluzioni di Big Data disponibili sul mercato come Hadoop, MapReduce e HBase.

Oltre al Data Scientist e al CDO possono essere identificati altri 4 ruoli per la costruzione di un team perfetto per la gestione dei sistemi BDA&BI (Ariker M. M. T., 2013).

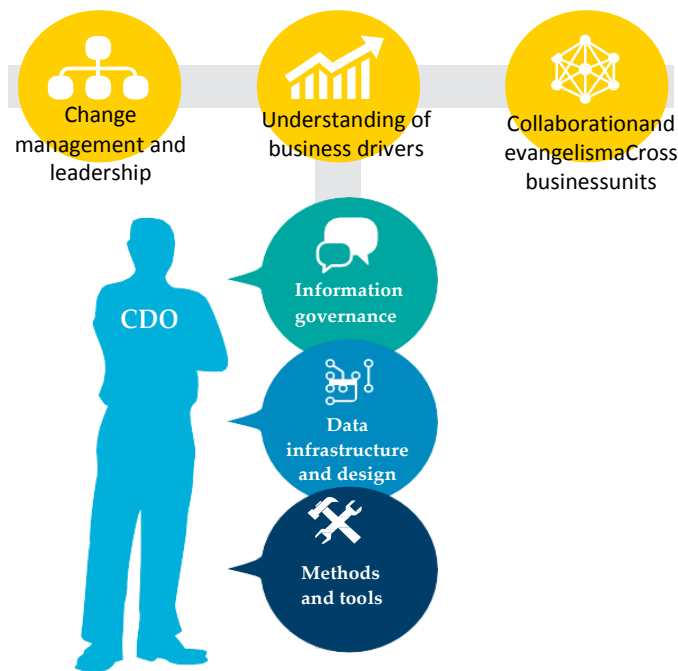


Figura 6: Set di competenze che deve avere il CDO per guidare il cambiamento, la diffusione di nuove idee e la collaborazione all'interno di tutta l'organizzazione (Van Rijmenam M., Chief Data Officer Profile, 2015)

- *Data Hygienists*: devono fare in modo che i dati siano puliti e accurati lungo tutto il loro ciclo di vita, per esempio devono verificare che i valori temporali presenti nei diversi Dataset siano gli stessi perché magari un Dataset considera i giorni in un anno solare (365), un altro solo i giorni lavorativi (260) e un altro ancora le ore in un anno (8675) ed eventualmente procedere alla correzione per poter fare dei confronti. Questa fase di pulizia deve essere svolta all'inizio dopo la raccolta dei dati e coinvolgere tutti i membri del team che li utilizzano.
- *Data Explorers*: mettono in ordine milioni di dati per identificare quelli realmente necessari perché, molto spesso, alcuni dati che in apparenza non hanno utilità da un punto di vista analitico non vengono memorizzati.
- *Business Solution Architects*: mettono insieme i dati scoperti e li strutturano in modo tale da poter essere interrogati in tempi adeguati da tutti gli utilizzatori.



- *Campaign Experts*: trasformano i modelli in risultati. Queste persone, che hanno una conoscenza completa e profonda sulla realizzazione delle campagne di Marketing, utilizzano le informazioni ricavate dai modelli per stabilire la priorità dei canali e mettere in sequenza le campagne; per esempio, basandosi sul comportamento storico di uno specifico segmento di clienti, identificano come soluzione migliore l'invio di un'e-mail seguita dopo 48 ore da una direct mail.

È molto importante mappare i movimenti di tutti i dati all'interno del team di Big Data e assicurarsi che tutto sia chiaro, in modo tale che ciascuna persona sia responsabilizzata su tutto il percorso e non solo sul suo compito specifico.

## **1.8 AZIONI E COMPORTAMENTI DA IMPLEMENTARE PER RIUSCIRE A SFRUTTARE I BIG DATA**

Approfondiamo ora una serie di comportamenti e azioni che le imprese possono mettere in atto per superare le barriere che ostacolano l'utilizzo dei Big Data e quindi la possibilità di ottenere benefici.

### **Focalizzazione sulla gestione del cambiamento**

La sfida cruciale che il management deve affrontare è il cambiamento del modo con cui prendere le decisioni. I leader delle organizzazioni devono innanzitutto selezionare poche aree, dalle quali possono trarre i maggiori benefici, su cui sperimentare i Big Data. Una nota azienda di telecomunicazioni si è focalizzata per esempio sull'applicazione di una soluzione analitica che permette di identificare in tempo reale il tasso di abbandono dei clienti al fine di migliorarne la gestione e una volta che il modello è stato sviluppato, è stata implementata una trasformazione delle prime linee di business, una riprogettazione dei flussi di customer-service e corsi di formazione sui nuovi strumenti per gli agenti (Court D., 2015).

Per creare valore le organizzazioni inoltre non devono partire da soluzioni analitiche complesse, ma devono procedere per passi. Nella prima fase i dati devono essere digitalizzati, strutturati e organizzati in modo tale da poter essere utilizzati direttamente, senza errori e con qualità garantita. In seguito è possibile applicare semplici analytics e solo in un secondo momento soluzioni analitiche avanzate, come algoritmi automatizzati e analisi in real-time.

### **Ridefinizione di ruoli e responsabilità e acquisizione di talenti**

Dato che una gran parte delle attività viene automatizzata e di conseguenza molti compiti eliminati, i manager devono procedere a una ridefinizione dei ruoli per riuscire a far leva sui Big

Data e per supportare il continuo sviluppo di nuove tecnologie. In una famosa assicurazione, in seguito all'introduzione di una soluzione analitica per la previsione della gravità dei reclami, i manager che prima si occupavano di questo hanno iniziato a processare solo i reclami eccezionali con un maggior livello di complessità e non più tutti (Court D., 2015).

Come sottolineato più volte, le organizzazioni hanno bisogno di talenti con profonde skill analitiche che siano in grado di fare analisi sui Big Data. È fondamentale che i leader organizzino il gruppo di talenti che hanno acquisito in modo tale da formare una comunità interna o un centro di eccellenza che sia connesso con il resto dell'organizzazione e che collabori con i leader del business.

I policy maker possono agire in diversi modi per favorire la fornitura di talenti. Innanzitutto possono promuovere iniziative educative negli ambiti della statistica e del decision science che aumentino il numero di laureati con tali competenze. In secondo luogo possono ridurre le barriere all'accesso a talenti che si trovano in altri Paesi promuovendo il lavoro a distanza e l'immigrazione e infine possono dare incentivi per la formazione dei manager e degli analisti sulle tecniche di Big Data (Manyika J., 2011).

### **Costruzione di una cultura orientata agli Analytics**

Senza una cultura orientata ai dati, tutti i benefici di cui abbiamo parlato in precedenza sono irraggiungibili. I leader all'interno delle organizzazioni devono avere una base di conoscenza delle tecniche analitiche in modo tale da essere in grado di utilizzare le nozioni estratte da analisi di questo tipo. A tal fine le aziende devono considerare queste competenze nei requisiti di assunzione e prevedere dei corsi di formazione che permettano ai leader di svilupparle. È fondamentale che i top manager siano consapevoli dell'impatto dei Big Data sulla gestione aziendale e in particolare dell'importanza dell'utilizzo dei marketing analytics e dell'approccio al marketing data-driven.

In UK il retailer Tesco ha sviluppato un approccio orientato ai dati sia per il top management che per le linee di business, integrando la customer intelligence in tutte le operations a tutti i livelli. Oggi i leader vengono fatti ruotare in diversi centri di analytics dove imparano gli elementi base relativi ai nuovi strumenti e al loro utilizzo per poi applicarli nei loro business: in questo modo essi non diventano né specialisti di analytics né data scientist ma iniziano a vedere applicazioni pratiche in casi reali e a capire le opportunità esistenti (Manyika J., 2011).

Le aziende hanno a disposizione un ampio set di opzioni per stimolare l'impegno dei dipendenti nell'apprendimento dei vari tool. Un'azienda leader nei servizi finanziari, per esempio, ha iniziato sviluppando una competizione che premia e riconosce quei team che hanno generato rilevanti insight attraverso strumenti analitici. In seguito ha stabilito la formazione in centri dove gli end-users imparano ad utilizzare gli strumenti self-service e ha promosso un programma di comunicazione per condividere l'entusiasmo generato in questo ambito tramite incontri, newsletter e comunicazione con i leader (Court D., 2015).

### **Utilizzo e integrazione di molti tipi di dati**

Per cogliere le nuove opportunità le aziende devono integrare dati provenienti da molteplici fonti, acquisendo l'accesso ai dati di terze parti e unendoli a quelli interni. In alcuni casi però le terze parti non vogliono condividerli perché li considerano un elemento chiave per ottenere un vantaggio competitivo oppure perché presentano degli errori che possono mettere a rischio la loro reputazione e quindi le aziende devono essere brave a fare proposte convincenti per ottenerne l'accesso.

### **Adozione di una strategia orientata all'integrazione di tecnologie**

Le aziende devono adottare una strategia di integrazione di tecnologie quali storage, computing e software analitici e di tecniche per memorizzare, organizzare e analizzare ampi dataset.

### **Garanzia della privacy e della sicurezza**

Le aziende devono garantire sia la privacy che la sicurezza, aspetti fondamentali visto che i dati viaggiano sempre di più entro e oltre i confini organizzativi. Per quanto riguarda la privacy non è sufficiente assicurare la conformità a leggi e regolamenti, ma le organizzazioni devono instaurare un rapporto di fiducia con clienti, partner di business e dipendenti e quindi comunicare loro in modo chiaro le politiche che intendono implementare, in modo che i vari stakeholder sappiano quali informazioni su di loro sono in circolazione e possono essere potenzialmente utilizzate.

Relativamente alla sicurezza invece, le aziende devono definire e implementare una strategia di rischio che includa la valutazione di tutti i rischi a cui sono sottoposte, dalla probabilità che gli hacker penetrino nei mainframe al rischio che le persone autorizzate utilizzino i dati a cui hanno accesso per scopi che vanno contro a quelli dell'organizzazione.

## Collaborazione tra CMO e CIO

Un'ulteriore che le aziende devono mettere in pratica *la collaborazione tra CMO e CIO* perché il primo ha a disposizione un enorme quantità di dati sui clienti da cui possono essere estratti insight, mentre il secondo ha le competenze necessarie per sviluppare le infrastrutture che permettono di creare le fondamenta dei Big Data e di trarre gli insight. I due devono lavorare insieme per riuscire a sfruttare i Big Data tuttavia, perché la collaborazione funzioni, è necessario che entrambi siano disposti a modificare il modo in cui lavorano. I primi devono capire che la tecnologia deve essere sfruttata e che il CIO è diventato un partner strategico da coinvolgere nella definizione della strategia di marketing.

La costruzione di questa relazione è abbastanza difficoltosa. Per sviluppare una partnership di successo il CMO innanzitutto deve definire gli obiettivi di business, i casi d'uso e i requisiti delle iniziative di Big Data e analytics, a questo punto il CIO può eseguire un'analisi dei costi e verificare la fattibilità del progetto. I due manager dovranno risolvere i trade-off tra tempo, costo e priorità. Per migliorare la comunicazione, soluzioni adottate da alcune aziende consistono nel posizionare i loro uffici sullo stesso piano o nell'organizzare cene ogni trimestre con lo scopo di costruire spirito di squadra e fiducia tra i due team. Un altro problema da considerare è la costruzione di un linguaggio condiviso dato che CMO e CIO parlano due lingue differenti (Ariker M., 2014).

Ci sono altre azioni da implementare per favorire la costruzione della loro relazione. Innanzitutto deve essere definito un framework relativo alla governance delle decisioni che indichi come CIO e CMO debbano lavorare insieme e supportarsi a vicenda.

In secondo luogo i due manager devono definire un'agenda comune per la definizione, la costruzione e l'acquisizione di competenze analitiche avanzate che porti alla costruzione di un Centro di Eccellenza dove lavorino persone sia del marketing che dell'IT.

Essi devono inoltre garantire la trasparenza dei processi: non solo devono definire con precisione i requisiti per l'utilizzo dei dati, ma devono incontrarsi regolarmente, ogni due settimane o ogni mese, per esaminare i progressi, identificare gli errori e mantenere alto l'impegno.

Dato che pochi CMO e CIO riescono a bilanciare in modo corretto business e tecnologia, è necessario assumere un "traduttore": il primo ha bisogno di assumere una persona che capisca le esigenze dei clienti e quelle di business ma che "parli geek", mentre al secondo serve una

persona con competenze tecniche che possieda però una base di conoscenza relativamente alle campagne di marketing e al business, come il Business Solution Architect che scopre i dati e li organizza al fine di analizzarli.

Infine, affinché la collaborazione abbia successo, è necessario identificare e focalizzarsi su un numero ridotto di piccoli progetti pilota per testare i team e i nuovi processi: questo approccio consente ai gruppi di sviluppare le best practice e di apprendere lezioni di valore che possono essere utilizzate per formare altri team.

## 1.9 ANALISI DELLE SURVEY

L'analisi di diverse survey riguardanti i temi che abbiamo affrontato, ci permette di capire quali siano effettivamente lo stato degli investimenti nei Big Data, i benefici che le aziende pensano di trarre dal loro utilizzo, le sfide e le barriere da superare per raggiungere tali vantaggi, le tipologie di dati e le tecnologie utilizzate.

La survey condotta nel 2014 da Gartner e rivolta a 302 IT e business leader di diversi settori (Servizi, Manufacturing, Banking, Pubblica Amministrazione, Retail, Sanità,...) evidenzia un interesse più concreto attorno ai Big Data, legato sia ad un aumento degli investimenti che ad un miglioramento delle offerte dei vendor. Il 40% delle aziende considerate infatti ha già investito, il 33% ha pianificato dei progetti entro i prossimi 24 mesi e solo il 24% non ha nessun programma relativo ai Big Data. Tuttavia solo il 13% delle organizzazioni ha un progetto in fase di sviluppo, il 27% sta affrontando l'implementazione di un progetto pilota e la sperimentazione, il 19% sta definendo la strategia e il 13% sta ancora raccogliendo informazioni e conoscenza. Questi dati mettono in luce come lo sviluppo dei progetti sia ancora in una fase arretrata.

Sono i servizi finanziari e la Sanità i settori che primeggiano nell'implementazione e nella pianificazione dei progetti Big Data e in particolare è il settore sanitario che intende svilupparli entro i prossimi 12 mesi (*Big Data and Analytics Survey condotta nel 2015 da IDG Enterprise, rivolta a 1139 IT decision maker di grandi, medie e piccole aziende in tutto il mondo e di svariati settori*).

Per quanto riguarda l'attitudine delle organizzazioni verso i Big Data Analytics, secondo il 46% questi rappresentano una nuova fonte di vantaggio competitivo e saranno fondamentali per il loro business, un altro 46% li considera un completamento del loro Data Warehouse e delle attività di Business Intelligence e solo l'8% li ritiene delle tecnologie utili da avere ma non la

priorità (*Wikibon Big DataAnalytics Adoption Survey 2014-2015 rivolta a 300 tra direttori e manager di aziende americane dei settori IT Technology Providers, Sanità, Manufacturing, Banking & Finance, Retail*).

Le aziende pensano di ottenere svariati benefici tramite lo sfruttamento dei Big Data: la survey condotta da IDG Enterprise mostra come una maggiore qualità del processo decisionale (61%) e il miglioramento della pianificazione e previsione (57%) siano i due principali vantaggi che le organizzazioni pensano di raggiungere. Dall'analisi di Gartner invece è emerso come il 68% delle aziende pensi di migliorare la customer experience, il 48% di rendere i processi più efficienti, il 45% di svolgere azioni di Marketing più personalizzate, il 42% di ridurre i costi e il 33% di sviluppare nuovi prodotti.

Per riuscire ad estrarre valore dai Big Data le aziende devono superare degli ostacoli, suddivisi in tecnologici e non tecnologici da WikiBon. Le aziende a cui è stato sottoposto il sondaggio, hanno indicato come prime quattro barriere tecnologiche la difficoltà di trasformare i dati in forme adatte per l'analisi, la difficoltà di integrare i Big Data con le infrastrutture esistenti, la difficoltà di fondere dati da disparate fonti e la mancanza di adeguate competenze. Le principali barriere non tecnologiche identificate sono invece la difficoltà nel rendere gli stakeholder d'accordo sulla definizione di Big Data, la difficoltà nel far comprendere il valore di questi agli utilizzatori finali, il problema della privacy e la complessità nel trasformare gli insight in azione.

Delle aziende indagate da IDG Enterprise il 48% ha individuato come sfida principale la mancanza di talenti nell'analisi e nella gestione dei dati e il 47% limiti di budget.

Per perseguire le opportunità offerte dai Big Data, deve essere esplorata una grande varietà di dati: i più citati sono i location-based Data (56%), seguiti dai dati testuali come email, slides, documenti Word e instant messages (48%). Vengono poi raccolti dati da social media (43%), immagini (39%), weblogs (37%), video (37%) e sensor Data (31%) (*Survey On The State Of Big DataAnalytics condotta nel 2015 da Forbes e rivolta a 316 dirigenti di grandi aziende globali*).

Alla domanda relativa al volume di dati che lo sviluppo di un progetto di BigDataAnalytics comporta all'interno delle loro organizzazioni, il 33% degli intervistati da WikiBon ha risposto tra i 50 e i 99 terabyte, il 28% tra 100 e 499 terabyte e solamente l'1% ha parlato di petabyte.

I BigData richiedono l'utilizzo di un mix di tecnologie nuove e vecchie per estrarre valore. Dalla survey di Gartner è emerso più della metà delle aziende (57%) utilizzano o pensano di usare principalmente il tradizionale enterprise Data warehouse, considerate la capacità di gestione, il

supporto analitico e le performance. Il 42% ha invece citato il cloud computing che, con le sue caratteristiche di elasticità e scalabilità, è in grado di supportare un progetto BigData, mentre per il 40% Hadoop è la soluzione ideale, data le sue capacità di memorizzazione ad hoc e di processamento di un'ampia varietà di dati. Il 34% ha invece indicato i Database relazionali, mentre il 20% i Database NoSQL.

Infine il 54% degli IT decision maker che hanno risposto alla survey di IDG Enterprise si aspetta che tali iniziative abbiano un impatto significativo o moderato entro i prossimi 12-18 mesi. Questo tuttavia è il pensiero delle grandi aziende, mentre per quelle piccole e medie l'impatto non sarà per niente rilevante.

## 2 L'IMPATTO DEI BIG DATA SUL MARKETING E SULLA CUSTOMER EXPERIENCE

Al giorno d'oggi i consumatori si aspettano che chi si occupa di Marketing nelle aziende conosca le loro preferenze, anticipi i loro bisogni e sfrutti tutte le informazioni a disposizione per costruire messaggi customizzati, passando dalla segmentazione dei clienti alla personalizzazione one-to-one in tempo reale, possibile grazie all'estrazione di insight a livello individuale. Il comportamento d'acquisto del consumatore è quindi cambiato negli ultimi anni e ha comportato il passaggio dall'approccio Marketing funnel, per cui le aziende pensano di spingere i clienti potenziali dalla consapevolezza all'acquisto puntando sulla generazione della maggior brand awareness possibile, a quello di consumer decision journey che riconosce un processo di acquisto molto più dinamico, in cui i consumatori subiscono influenze in svariati momenti (Bhandari R., 2014). Questi ultimi hanno acquisito nel tempo sempre più potere, portandoli ad occupare una posizione di vantaggio rispetto alle aziende: essi sono infatti liberi di cercare offerte low-cost per i loro prodotti preferiti, di utilizzare filtri antispam per evitare di ricevere promozioni e comunicazioni di Marketing indesiderate. Il compito delle organizzazioni di soddisfare i propri clienti diventa quindi molto più difficile (Tarka P., 2014). Una leva a disposizione delle organizzazioni per riuscire in questa missione sono i dati ed è proprio per questo motivo che dedichiamo questo secondo capitolo all'impatto dei BigData sul Marketing.

### 2.1 MARKETING ANALYTICS

I *MarketingAnalytics* sono un approccio abilitato dalle tecnologie e supportato da modelli che sfrutta i dati di mercato e dei clienti per migliorare il processo di decision making in ambito Marketing (Germann F., 2012). Queste soluzioni sono molto utili in un contesto come quello odierno, caratterizzato da elevata incertezza dovuta alle variazioni continue delle preferenze dei clienti, relativamente alle caratteristiche dei prodotti, dei prezzi e dei canali, per cui per un'azienda diventa molto più complicato prendere decisioni. I *MarketingAnalytics* offrono un aiuto in quanto permettono di monitorare gli impulsi del mercato e di fornire avvertimenti sui cambiamenti dei gusti dei consumatori.

L'obiettivo è sviluppare un approccio *Data-driven Marketing*, che consiste nella raccolta e nella combinazione di dati online e offline per fornire rapidamente informazioni sui clienti provenienti da diversi canali integrati tra di loro, in modo da realizzare campagne di Marketing personalizzate e soddisfare i loro bisogni (Teradata, 2015). Non sorprende quindi che sia il



Marketing più di tutte le altre funzioni aziendali a richiedere l'utilizzo dei BigData. Questo cambiamento mette inoltre in evidenza il suo ruolo strategico, che porta le aziende ad agire proattivamente, anticipando la domanda per generare una capacità di risposta in real-time.



*Figura 7: Fattori alla base della necessità dei Marketing Analytics*

### 2.1.1 TIPOLOGIE DI DATI RACCOLTI

Per definire le strategie di Marketing con l'obiettivo di creare messaggi personalizzati e acquisire nuovi clienti, devono essere raccolti dati incentrati sui consumatori. Si tratta di dati anagrafici, dati provenienti dalle carte fedeltà, dati relativi alla customer satisfaction, dati di customer service, dati sul comportamento dei clienti, dati di acquisto e delle preferenze di acquisto, dati riguardanti l'engagement del consumatore e dati di Digital Marketing(Teradata, 2015). Il loro ammontare è enorme in quanto non ci sono solo dati raccolti e memorizzati dalle organizzazioni nel CRM transazionale, in quello collaborativo e nei sistemi gestionali, ma anche dati provenienti da mobile, Web, mail, call center, social network, Direct Marketing e agenti. Questi vengono creati tramite i click sulle pagine Web, i cookie e i commenti postati su prodotti e servizi. La sfida consiste nell'incrociare dati strutturati e non, interni ed esterni in modo da estrarre insight significativi.

### 2.1.2 DALL'APPROCCIO TRADIZIONALE ALL'APPROCCIO PREDITTIVO

Nel contesto odierno descritto, il modello classico di analisi dei dati di Marketing, che si basa sulla storia passata senza fare predizioni e utilizza fogli di calcolo ma nessuno strumento

diagnostico, è inadeguato. Innanzitutto c'è un ritardo tra l'analisi dei dati secondari e gli eventi effettivi: la tipica valutazione semestrale del Marketing mix non è sufficiente considerata la competizione crescente. In secondo luogo le analisi di Marketing vengono affidate a vendor esterni e prima che vengano inviate alle organizzazioni può passare molto tempo, impedendo loro di agire tempestivamente. Infine questo approccio non considera i nuovi canali di comunicazione molto attrattivi (Tarka P., 2014).

Per superare questi limiti le aziende sono passate a modelli predittivi che permettono di rispondere tempestivamente alle esigenze del mercato, basandosi sui seguenti elementi:

- integrazione di dati provenienti da fonti interne ed esterne che vengono analizzati con un'ampia gamma di tecniche di visualizzazione;
- utilizzo di analisi di scenario what-if per simulare cosa potrà accadere nel futuro nel caso di cambiamento di una variabile decisionale;
- passaggio da un progetto singolo a un processo continuo in cui gli analisti fanno simulazioni e previsioni, per determinare l'impatto delle variazioni della spesa in advertising e in Marketing per diversi prodotti, mercati e canali;
- analisi condotte internamente.

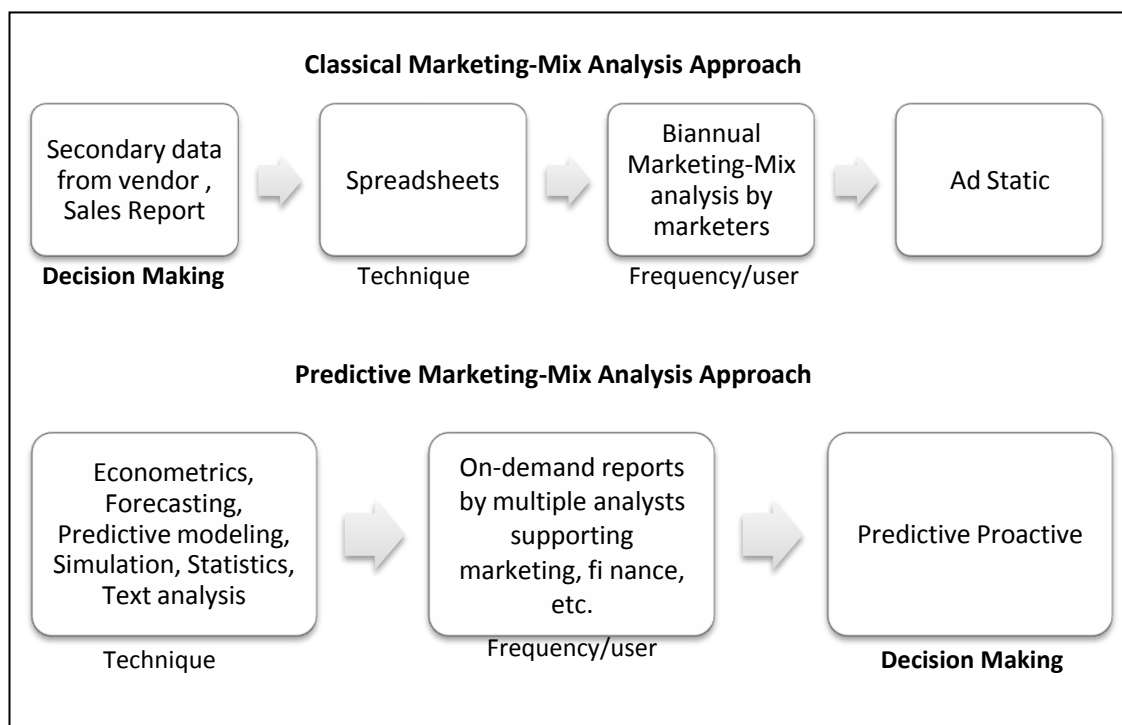


Figura 8: Approccio classico e predittivo all'analisi del Marketing mix (Tarka P., 2014)

### 2.1.3 BENEFICI CONSEGUIBILI DALL'UTILIZZO DEI BIG DATA NEL MARKETING

La raccolta dei BigData e l'utilizzo delle soluzioni di MarketingAnalytics permettono innanzitutto di *migliorare il processo di decision making*: viene esplorato un maggior numero di opzioni e valutato l'impatto delle diverse variabili decisionali, di conseguenza le decisioni prese sono più accurate e consistenti (Germann F., 2012).

In secondo luogo le organizzazioni possono individuare *nuove opportunità*, canali o mercati, sviluppando quindi modelli di business e strategie innovative ed efficaci.

L'elemento centrale dell'utilizzo dei Big Data in ambito Marketing è però la *visione a 360° del cliente*, che si ottiene andando ad integrare tutti i dati che le aziende possiedono con i second e i third party Data(IBM). La conoscenza approfondita del consumatore consente di:

- *Creare soluzioni di Marketing customizzate*, prevedendo quindi per ciascuno i migliori metodi di comunicazione, i canali, i messaggi e i tempi di invio. Personalizzazione significa estrarre insight relativamente al singolo consumatore: l'obiettivo delle aziende è infatti quello di sviluppare un approccio di Marketing individuale (Teradata, 2015). Inoltre dato che oggi i consumatori utilizzano social media, svariati dispositivi mobile e dispongono di molte informazioni, le organizzazioni devono interagire con loro attraverso i diversi canali in tempo reale, perseguendo perciò un approccio di multicanalità.
- *Inviare offerte di Cross-sell e up sell real-time*.
- *Ridurre i costi delle campagne Marketing*, rivolgendole ai consumatori che risponderanno con maggiore probabilità (Tarka P., 2014).
- *Anticipare i desideri e le esigenze dei clienti*, in modo da generare una capacità di risposta in real time, favorendo quindi un processo decisionale molto più rapido.

Fornire l'offerta giusta al momento giusto ai clienti che ne hanno bisogno impatta positivamente sulla loro *customer experience*, aumentando quindi il loro livello di soddisfazione. Le conseguenze sono sia *l'incremento del tasso di mantenimento della clientela* e quindi la trasformazione di clienti occasionali in clienti abituali sia *l'acquisizione di nuovi clienti*.

Il risultato finale è *l'aumento della redditività* dell'azienda e un maggior ritorno sull'investimento di Marketing (MROI).

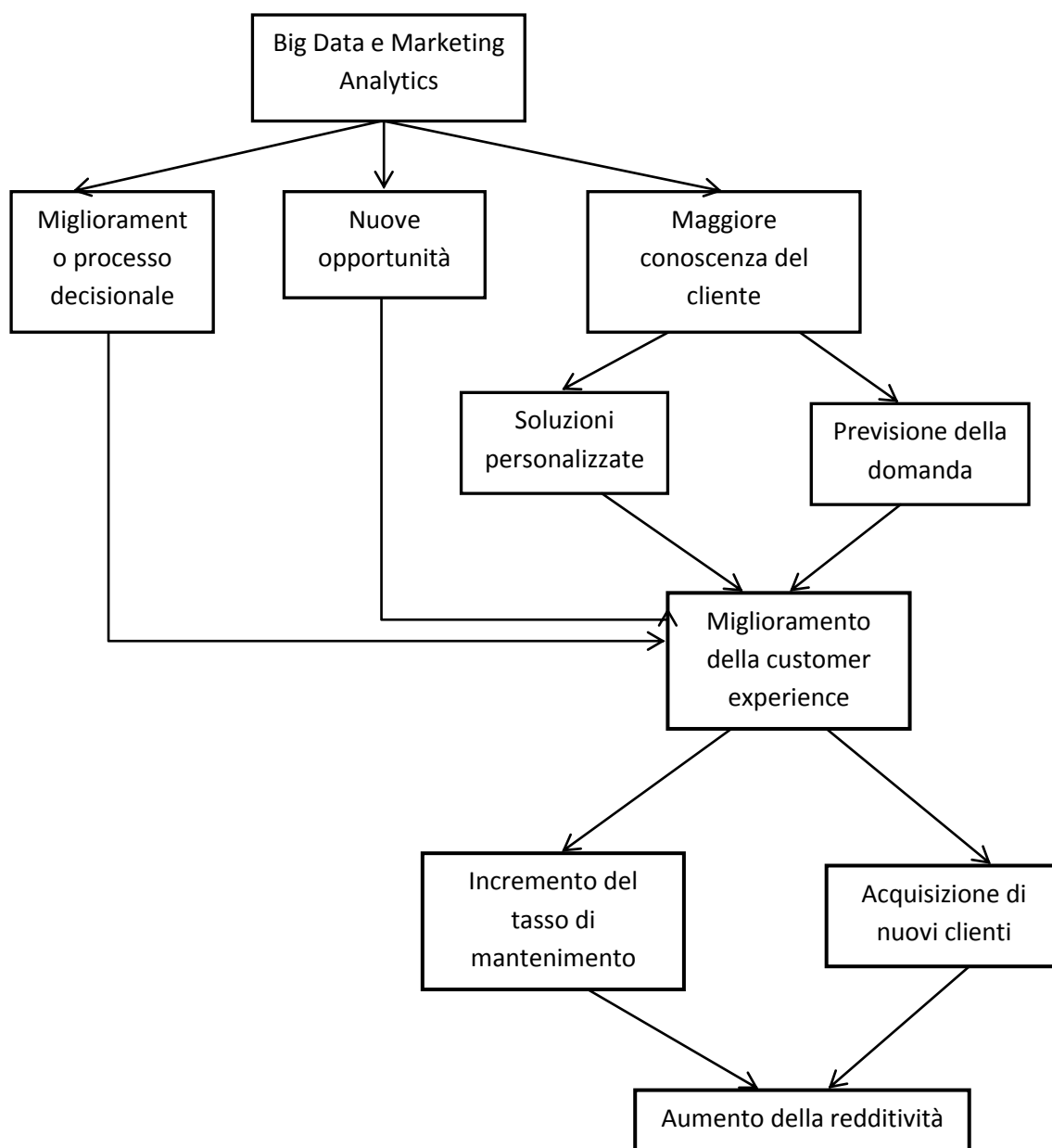


Grafico 3: Benefici conseguibili dall'utilizzo dei Big Data nel Marketing

#### 2.1.4 PASSI DA SEGUIRE PER RIUSCIRE A SFRUTTARE I MARKETING ANALYTICS

Le organizzazioni devono seguire tre fasi per sfruttare a pieno le potenzialità dei MarketingAnalytics e riuscire a trarne dei vantaggi (Bhandari R., 2014).

##### 1. Individuare il miglior approccio analitico

I team hanno a disposizione diversi strumenti e metodi dei quali devono valutare pro e contro per individuare quello migliore ai fini dei loro obiettivi strategici. In tabella riportiamo quelli principali.

Approccio	Descrizione	Pro	Contro
Marketing-mix modeling (MMM)	Approccio di advanced Analytics che utilizza i Big Data per determinare l'efficacia della spesa nei diversi canali. Questo collega statisticamente gli investimenti in Marketing ai driver di vendita, includendo variabili esterne come la stagionalità e le attività dei competitor al fine di scoprire effetti come i cambiamenti degli individui e dei segmenti nel tempo o le differenze tra le attività offline, online e dei social media.	È adatto sia per il raggiungimento di obiettivi strategici di lungo termine che per la pianificazione tattica	Richiede dati di alta qualità sulle vendite e sulla spesa di Marketing relativi a molti anni. Non è in grado di misurare attività che variano poco nel tempo e gli effetti di lungo termine di un investimento in un nuovo touchpoint. Inoltre è necessaria una profonda conoscenza di econometria per comprendere il modello e tool di scenario-planning per stabilire le implicazioni sul budget delle diverse decisioni di investimento.
Reach, cost, quality (RCQ)	Approccio euristico che disaggrega ciascun punto di contatto con il cliente nelle sue componenti, tra cui il numero di consumatori raggiunti e la qualità	È semplice da utilizzare e supera i limiti dell'approccio MMM: può essere usato quando i dati sono limitati, quando il tasso di spesa è costante durante	Non tiene in considerazione gli effetti della rete e delle interazioni e dipende molto dalle ipotesi alla sua base.

	dell'engagement, utilizzando sia dati strutturati che non.	l'anno e quando le cui componenti sono difficili da isolare.	
Attribution Modeling	Insieme di regole e algoritmi che gestiscono come assegnare il credito a ciascun touchpoint online per convertire le visite in vendite. In particolare vengono usati modelli statistici, tecniche di regressione e algoritmi di bidding.	È all'avanguardia rispetto agli altri due approcci e permette di valutare il successo dell'investimento online.	Ha a disposizione una base di dati limitata perché dipende dai cookie e quindi è difficile capire l'importanza di ciascun touchpoint.

Tabella 5: Approcci analitici per lo sfruttamento dei Marketing Analytics

## 2. Integrare le competenze per generare insight

Per scoprire insight significativi e ottenere migliori risultati sia nel breve che nel lungo termine è necessario implementare un approccio integrato, ovvero utilizzare due o più tra gli strumenti appena descritti.

## 3. Porre l'approccio analitico al centro dell'organizzazione

Gli uomini di Marketing devono collaborare con Data scientist, analisti digitali e "translators", persone che comprendono gli Analytics e allo stesso tempo parlano il linguaggio del business, per formulare ipotesi, mettere in discussione le assunzioni e raggiungere gli obiettivi prefissati.

Gli insight estratti dal consumer decision journey e dall'allocazione del Marketing-mix dovrebbero essere utilizzati per scegliere quali media utilizzare e i risultati effettivi dovrebbero essere confrontati con quelli previsti per aggiustare il budget da utilizzare.

### 2.1.5 OSTACOLI DA SUPERARE PER SFRUTTARE I MARKETING ANALYTICS

Aspetto molto importante per lo sviluppo e lo sfruttamento dei MarketingAnalytics è l'impegno del team di top management che non solo deve garantire la presenza di risorse necessarie quali skill analitiche, dati e tecnologie adeguate, ma deve anche promuovere una cultura che supporti l'utilizzo di queste soluzioni (Germann F., 2012). Tuttavia questi rappresentano degli aspetti critici che la maggior parte delle aziende si trovano ad affrontare, insieme ad altri. In particolare possiamo distinguere 4 diverse tipologie di barriere da superare, come indicato nel seguente grafico: tecniche, culturali, organizzative e legate alle competenze.

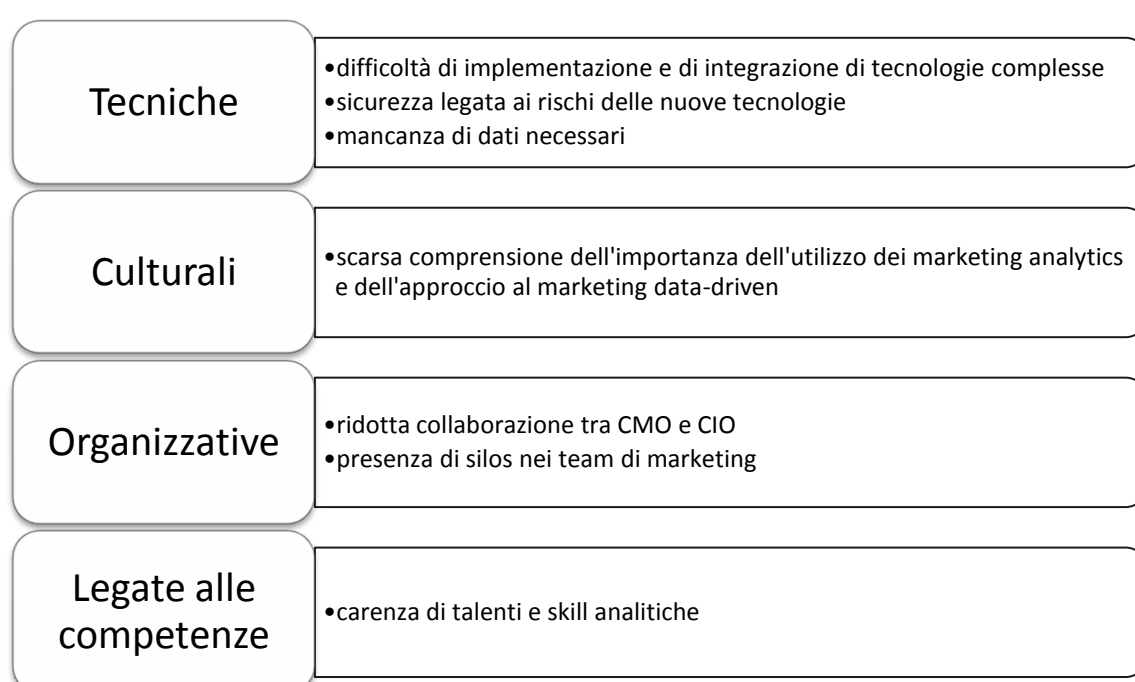


Grafico 4: Barriere allo sfruttamento dei Marketing Analytics

Le aziende devono innanzitutto imparare a gestire la complessità delle tecnologie e dei task e la questione relativa alla sicurezza, considerato l'incremento dei rischi legati alle nuove soluzioni e la paura e la cattiva informazione in quest'area. È quindi utile che chi si occupa di Marketing sia aggiornato su questi temi. Inoltre, mentre la funzione Marketing è conscia dei benefici che l'utilizzo dei DataAnalytics comporta, i dirigenti delle altre funzioni non lo sono e questa mancanza di consapevolezza rappresenta un freno alla loro adozione. Non meno importante è la ridotta collaborazione tra gli uomini IT e quelli di Marketing: per rendere il MarketingData-driven serve cooperazione tra le due funzioni, che devono condividere i dati in modo da migliorare il processo di decision making. I cosiddetti silos sono presenti anche all'interno della

funzione Marketing stessa: i team agiscono come entità a se stanti e questa mancata interazione porta i gruppi a non avere una visione unica delle campagne e dei clienti attraverso i diversi canali, con effetti negativi sul servizio offerto ai consumatori. Altra barriera è l'assenza di talenti, di persone con competenze adeguate che siano in grado di gestire i MarketingAnalytics.

## 2.2 SOCIAL ANALYTICS

Elemento molto importante per il Marketing è rappresentato dai tools di *social Analytics* che monitorano, analizzano, misurano ed interpretano le interazioni digitali e le relazioni tra persone, topic, idee e contenuti sui social. Questi fanno leva su svariate tecniche, tra cui la sentiment analysis, il NLP e la network analysis, la text analysis, i modelli predittivi e la classificazione di topic, persone e contenuti. L'attenzione delle aziende verso questi strumenti è in continua crescita, considerati i grandi volumi dei commenti sui prodotti postati ogni giorno su social media e il potenziale offerto per sviluppare efficaci strategie di Marketing, di customer relationship management e di progettazione del prodotto (Lau R., 2014).



Figura 9: Social Media (Google, 2014)



### 2.2.1 LIVELLI DI MATURITÀ DELLE INIZIATIVE DI SOCIAL ANALYTICS

È possibile distinguere cinque diversi livelli di sviluppo crescente dei Social Analytics, a seconda del grado di maturità analitica e tecnologica dei metodi utilizzati dalle organizzazioni (Osservatorio Big Data Analytics & Business Intelligence, 2014).

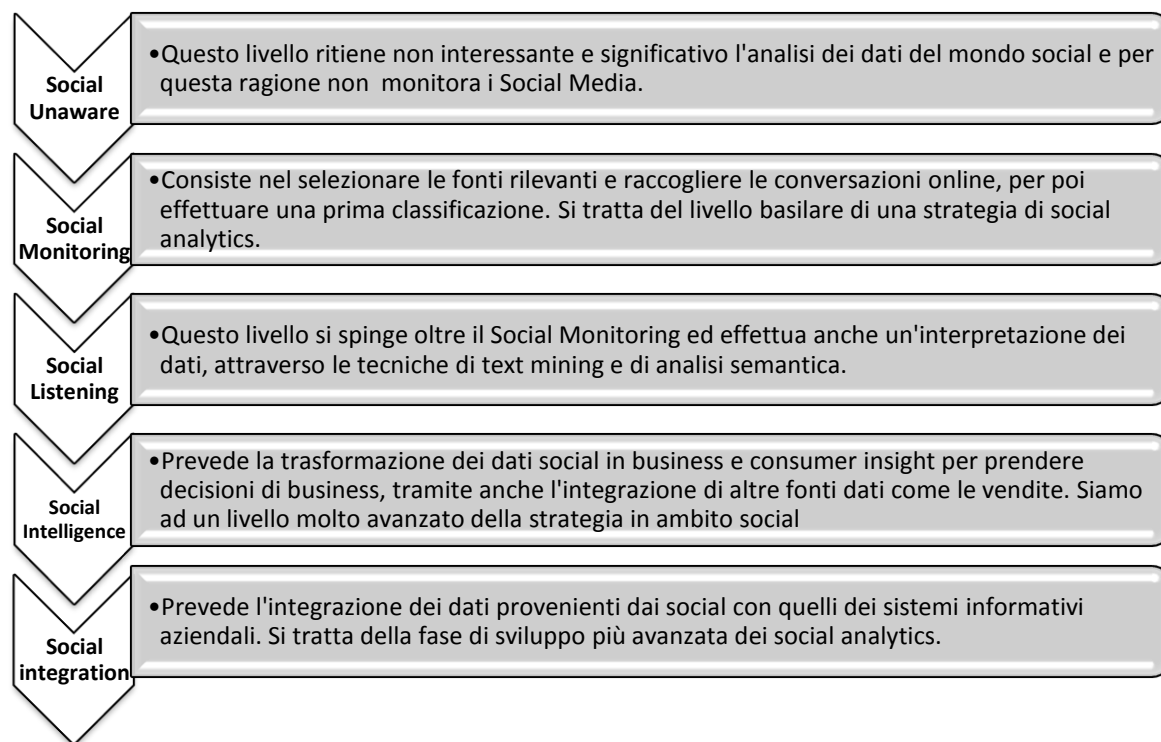


Grafico 5: Livelli di maturità delle iniziative di Social Analytics

Per massimizzare il Social Media ROI le aziende devono arrivare almeno al quarto step della scala appena descritta: non basta rilevare le metriche e visualizzare i risultati, ma è necessario stabilire quali metriche misurare in base agli obiettivi social prefissati e interpretare i risultati per estrarre insight rilevanti per il processo decisionale (BIT BANG, 2014). Le aziende devono costruire a tal fine un Framework che prevede appunto di:

- Specificare i traguardi che si vuole ottenere dalla Social Intelligence, che devono essere S.M.A.R.T , ovvero specifici (specific), misurabili (measurable), conseguibili (attainable), rilevanti (reliable) e puntuali (timely).
- Definire gli obiettivi social, che rappresentano i mezzi intermedi che consentono di raggiungere i risultati finali. I principali sono la crescita di reach, lo sviluppo del dialogo con e tra utenti, la generazione di interazioni e la promozione dell'advocacy.

- Selezionare le metriche e i KPI rilevanti per la valutazione delle performance della strategia perseguita e per guadagnare insight significativi. Un esempio è lo Share of Voice, percentuale di conversazioni riguardanti un'azienda sul totale di conversazioni relative al settore, che aiuta a capire il livello di influenza di un brand sul Web. Un altro è l'Engagement, calcolato come somma delle user interactions (Like, Commenti, Share, Retweet, etc), può essere valutato in senso assoluto, come percentuale di utenti che interagiscono sul totale dell'audience (Engagement Rate) o come percentuale di engagement generata da contenuti pubblicati dall'azienda rispetto all'engagement generato da tutti i contenuti pubblicati dai brand del settore (Share of Engagement). Questa metrica misura quanto i contenuti pubblicati stimolino l'interazione da parte dell'audience e quanto la propria user base venga coinvolta adeguatamente (BIT BANG, 2015).

Business Objective	Key Performance Indicator	Vendors to Watch
Foster Dialog	Share of Voice	Alterian SM2, Radian6, Scout Labs, Statsit, Trendrr, Visible Technologies
	Audience Engagement	Coremetrics, Webtrends, Radian6, Scout Labs, Converseon, Filtrbox (Jive), Visible Technologies
	Conversation Reach	Alterian SM2, Radian6, Scout Labs, Social Radar, Statsit, SWIX, Trendrr, Visible Technologies
Promote Advocacy	Active Advocates	Biz360, Filtrbox (Jive), Radian6
	Advocate Influence	Cymfony, Filtrbox (Jive), Lithium, Radian6, Razorfish (SiM Score), SAS, Telligent, Twitalyzer, Visible Technologies
	Advocacy Impact	Coremetrics, Lithium, Omniture, Webtrends, SWIX, Telligent
Facilitate Support	Resolution Rate	Filtrbox (Jive), RightNow Technologies, Salesforce.com, Telligent
	Resolution Time	Filtrbox (Jive), RightNow Technologies, Salesforce.com, Telligent
	Satisfaction Score	ForeSee Results, iPerceptions, Kampyle, OpinionLab
Spur Innovation	Topic Trends	Alterian SM2, Cymfony, Filtrbox (Jive), Radian6, SAS, Scout Labs, Social Mention, Social Radar, Trendrr, Visible Technologies
	Sentiment Ratio	Alterian SM2, Converseon, Cymfony, Filtrbox (Jive), Radian6, SAS, Scout Labs, Social Radar, Trendrr, Visible Technologies
	Idea Impact	Biz360, Cymfony, Filtrbox (Jive), Lugtron, Radian6, Scout Labs, Visible Technologies

Tabella 6: Social Media Intelligence Framework (BIT BANG, 2014)

## 2.2.2 INIZIATIVE DI SOCIAL ANALYTICS

I Social Analytics possono essere utilizzati dalle aziende per svariate applicazioni, che possono portare ad un vantaggio competitivo sui concorrenti. In tabella riportiamo quelle principali (BIT BANG, 2015), (BIT BANG, 2015), (BIT BANG, 2014).

Iniziativa	Descrizione	Strategia da seguire
Valutare la credibilità del brand	Tramite i social media per le aziende è possibile osservare il comportamento del consumatore e le conversazioni in modo continuo nel tempo, riuscendo quindi a individuare le opinioni dei clienti.	<ul style="list-style-type: none"> <li>• Impostare una strategia di Social Listening mirata alla raccolta di termini legati all'immagine dell'azienda, come la soddisfazione dei dipendenti e la qualità del servizio clienti.</li> <li>• Individuare una lista di temi problematici relativi all'azienda, ai quali gli utenti sono più interessati e una di temi riguardanti i competitor che stanno ricevendo opinioni positive.</li> <li>• Confrontare le due liste per valutare iniziative e modifiche di business al fine di aumentare la credibilità del brand.</li> </ul>
Identificare i bisogni dei clienti	Le organizzazioni possono ascoltare le comunicazioni pubblicate dai consumatori sui social relativamente ai prodotti, alle loro caratteristiche e alle innovazioni che desiderano vedere sul mercato.	<ul style="list-style-type: none"> <li>• Individuare i profili social pubblici dei propri clienti attraverso per esempio email-matching e login tramite le proprie properties.</li> <li>• A partire dalle informazioni disponibili come quelle del profilo e conversazioni sui prodotti, hobby e interessi, costruire un identikit di</li> </ul>

ciascun cliente.

- Eseguire il report dei risultati e diffonderlo in tutte i dipartimenti dell'azienda.

Caratterizzare il brand

Dato l'aumento del potere del consumatore, il brand oggi non viene più definito dal dipartimento di Marketing dalle aziende, ma dalle opinioni espresse dagli utenti sul Web e anche in questo caso alle organizzazioni non resta che ascoltarle.

- Integrare i dati social dei profili dei clienti e la visione interna dell'azienda, per definire nuovamente il brand.
- Ricorrere alla Social Intelligence per formulare strategie che attirino e fidelizzino i consumatori.

Ottimizzare la comunicazione

Tramite i dati degli ambienti social, le organizzazioni possono estrarre insight per capire come comunicare ed interagire con il cliente in modo efficace, individuando i contenuti e il linguaggio da utilizzare.

- Raccogliere i contenuti riguardanti brand e competitor derivanti dalle attività sui social.
- Tramite text Analytics analizzare i testi dei contenuti e individuare le keyword utilizzate per definire brand e competitor e concentrarsi su quelle affini agli interessi dell'audience.
- Attraverso le metriche di social Analytics, quali like e retweet, identificare le tipologie di contenuti che interessano maggiormente

gli utenti.

- Utilizzare tutti questi dati per definire la strategia di comunicazione migliore.

#### Individuare gli influencer

Le aziende riescono a identificare persone capaci di influenzare i pensieri e le decisioni degli altri grazie a ciò che dicono e scrivono. Ci sono tre elementi che costituiscono la social influence: *rilevanza* (affinità tra l'individuo e il settore in cui opera la compagnia), *portata* (quantità di individui raggiunti dai contenuti dell'influencer) e *risonanza* (frequenza di pubblicazione, engagement ottenuto e livello di diffusione ad opera del pubblico dei contenuti dell'influencer).

- Monitorare le conversazioni sull'argomento di interesse per l'azienda in modo da identificare che gli utenti sotto osservazione stiano effettivamente parlando di questo.
- Ordinare gli utenti rispetto ad indici di frequenza, come metriche tipiche dei social media (amici per Facebook, follower per Twitter, subscriber per Youtube) o indici più complessi come il Klout che include fattori provenienti da varie presenze web del medesimo utente.
- Verificare in modo qualitativo che gli individui identificati possano effettivamente essere considerati influencers per il settore.

#### Monitorare i competitor

Sfruttando i dati dai social media, le organizzazioni sono in grado di monitorare le performance e di tracciare le strategie e i risultati

- Tramite benchmarking, osservare le performance social dei competitor per contestualizzare i propri

dei loro concorrenti.	<p>risultati in modo analitico.</p> <ul style="list-style-type: none"> <li>• Utilizzare il Social Listening delle conversazioni che riguardano i competitor, per individuare eventuali loro punti di forza o di debolezza.</li> <li>• Analizzare la performance del piano editoriale del concorrente per identificare i contenuti che hanno avuto maggior successo con l'audience.</li> </ul>
-----------------------	---

Tabella 7: Iniziative di Social Analytics

### 2.2.3 SFIDE DA AFFRONTARE PER SFRUTTARE I SOCIAL ANALYTICS

Ci sono alcune sfide che le aziende devono affrontare per progettare efficaci tools di social Analytics. In primo luogo identificare i prodotti e le loro caratteristiche dai commenti online è complesso essendo questi non strutturati e scritti in forma libera. In secondo luogo le opinioni e le emozioni dipendono dal contesto, di conseguenza non è sufficiente applicare il dizionario del sentiment per estrarre social intelligence dai commenti dei consumatori. Infine collegare i sentimenti alle caratteristiche specifiche dei prodotti è abbastanza complicato in quanto queste ultime nei commenti degli utenti sono molto spesso riferite ad aspetti più generali del prodotto sia impliciti che espliciti (Lau R., 2014).

Questi problemi possono però essere risolti ricorrendo ad una metodologia di social Analytics sostenuta da un algoritmo di fuzzy product ontology che permette di svolgere un'analisi sensibile al contesto e orientata ai sentimenti legati agli aspetti del prodotto, in grado di gestire l'incertezza. Essa prevede 7 step:

1. *Query Processor*: l'utente seleziona una categoria di prodotto o un prodotto specifico per il quale vuole estrarre social intelligence dai commenti online.
2. *Social Comments Retrieval*: i commenti dei consumatori vengono recuperati attraverso gli Web Services o le API fornite dai motori di ricerca o dai social media.

3. *Social Comments Crawler*: un software dedicato recupera le informazioni sui prodotti, sulle loro caratteristiche e i relativi commenti.
4. *Text Pre-processor*: i commenti e le descrizioni dei prodotti vengono preprocessati: le caratteristiche vengono rappresentate da frasi nominali, mentre i sentimenti da aggettivi o avverbi.
5. *Product Ontology Miner*: le relazioni tra le caratteristiche dei prodotti, i loro aspetti e i sentimenti associati vengono identificate attraverso un algoritmo di product ontology mining e una sentiment analysis orientata agli aspetti dei prodotti. I risultati vengono quindi visualizzati per mezzo di un programma grafico open source.
6. *Aspect Oriented Sentiment Analyzer*: l'analizzatore dei sentiment determina la polarità positiva, negativa o neutrale per ciascuna coppia aspetto-sentimento rilevata dai commenti dei consumatori.
7. *Aspect-oriented Product Recommender*: ogni coppia viene poi analizzata e gli viene assegnato un punteggio sulla base dei commenti recuperati. Quindi i prodotti appartenenti ad una stessa categoria possono essere confrontati sulla base di questi punteggi.

#### 2.2.4 TENDENZE FUTURE

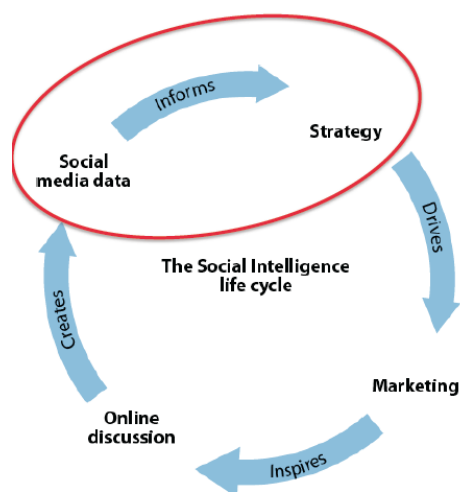
Vediamo ora le tendenze più recenti relative ai Social Analytics e verso quali direzioni si muoveranno nel futuro.

- ✓ Passi avanti nel calcolo del Social Media ROI grazie all'evoluzione degli strumenti di misurazione digitale e all'abbattimento dei Data Silos e nella definizione dei KPI che diventeranno sempre più sofisticati e si concentreranno principalmente sull'engagement piuttosto che sulla portata di messaggi pubblicitari.
- ✓ Miglioramento nella profilazione dell'utente attraverso l'utilizzo di metodi che considerano maggiormente aspetti quali demographics, interessi, hobbies, caratteristiche lavorative, etc per implementare strategie di Ad Targeting fortemente personalizzate. Sempre nell'ambito della profilazione un nuovo obiettivo è il collegamento dei vari account social e digital al singolo individuo a cui fanno capo.
- ✓ Maggiore disponibilità per le aziende di insight da cui trarre business value. Facebook per esempio ha attivato il servizio TopicData, che consente di ottenere insight dall'intero firehose del social network, mentre prima il servizio di Graph Search permetteva solamente la raccolta di post con privacy pubblica e quindi i contenuti accessibili erano inferiori. L'accesso è esclusiva di Datasift e molte aziende specializzate

in Social Analytics hanno deciso di collaborare con essa per sfruttare queste potenzialità, fornendo quindi diversi vantaggi ai clienti (BIT BANG, 2014).

- ✓ Possibilità di misurazione delle prestazioni del proprio Social CRM grazie all'integrazione tra le piattaforme di Social Monitoring e quelle di CRM.
- ✓ Nuove frontiere per il Social Listening: da un lato si pensa di poter sfruttare gli strumenti tradizionali per individuare situazioni specifiche quali l'intenzione d'acquisto, le richieste di nuovi prodotti e feature o le lamentele sui prodotti e di agire quindi tempestivamente (BIT BANG, 2014). Dall'altro lato si prevede verranno sviluppati algoritmi di sentiment più precisi in grado di cogliere stati d'animo più complessi come rabbia, noia e confusione e superare quindi la classica distinzione tra "positivo/negativo e neutrale" (BIT BANG, 2015). Un nuovo trend è rappresentato dal Visual Listening, che consiste nell'analisi delle immagini, fondamentale dato che oggi sui vari social vengono pubblicate 500 milioni di foto al giorno e sono proprio i contenuti visuali ad attrarre maggiormente i consumatori. È possibile individuare quando e dove il logo e altri riferimenti visivi al brand vengano utilizzati ed estrarre quindi insight su come i propri prodotti e il proprio brand vengano associati a contesti e ad attività. Lo sviluppo di questa nuova tendenza è favorita dall'evoluzione di algoritmi di analisi delle immagini sempre più sviluppati, che possono essere utilizzati dalle aziende.
- ✓ Analisi delle immagini, non solo per determinare il sentiment di un post, come viene già fatto, ma identificare come queste vengano utilizzate in associazione al logo per trarne
- ✓ insight significativi (BIT BANG, 2015).

## 2.2.5 IMPLICAZIONI SUL MARKETING



L'impatto della social intelligence sul Marketing può essere rappresentato da un loop continuo, mostrato in figura. Essa informa la strategia che guida il Marketing, il quale ispira a sua volta le conversazioni online che generano i dati social. Quindi l'analisi e la gestione dei dati dei clienti provenienti dai social vengono sfruttati per definire e rivedere i programmi di Marketing e di business.

Figura 10: Impatto della Social Intelligence sul Marketing (Bellini L., 2012)



## 2.3 PRINCIPALI AMBITI DEI PROGETTI DI BIG DATA NEL MARKETING

Lo sfruttamento dei Big Data in ambito Marketing rappresenta un enorme potenziale, tanto che le aziende si stanno dedicando e hanno un grande interesse verso progetti che prevedono il loro utilizzo in quest'area. Oltre ai *Social Analytics* già affrontati, altri sono il *Direct* e il *Digital Marketing*, la *Customer Micro-segmentation*, il *Location-based Marketing*, l'*In-store Analysis* e il *Cross-Selling/Up-Selling*.

### 2.3.1 DIRECT E DIGITAL MARKETING

Il Direct Marketing comprende tutte le tecniche di Marketing che consentono alle aziende di comunicare in modo mirato e personalizzato direttamente con il cliente o l'utente finale. La continua e significativa crescita di internet e della sua importanza ha comportato il rapido sviluppo del *Digital Marketing*, che assume la forma di display advertising, contenuti su Facebook, video clip su Youtube, e-mail personalizzate e molto altro. Le aziende per fare Digital Marketing oggi possono contare sull'enorme ammontare di informazioni degli utenti, che trascorrono ore e ore al giorno su Internet, relative ai loro interessi, ai contenuti delle loro comunicazioni, agli acquisti che fanno e molto altro (Hazan E., 2013).

Il Direct Marketing si serve di molte tecniche di Big Data, oltre che per identificare i clienti più profittevoli e quelli che risponderanno con maggiore probabilità, soprattutto per profilare i clienti, in modo da prevedere anche il comportamento di quelli sconosciuti. Vengono utilizzate sia tecniche di apprendimento supervisionato, come i modelli di ottimizzazione, le reti neurali bayesiane e gli alberi decisionali sia quelle non supervisionate, tra cui il clustering. Per ottenere risultati migliori l'ideale è combinare diverse tecniche. Un approccio consiste per esempio nell'utilizzare le tecniche di self-organizing maps (SOM) e di estrazione delle dimensioni salienti (SD): la prima consente di rappresentare un Dataset contenente molti input con una mappa di dimensioni minori per esplorare i dati e utilizzare tecniche come il clustering in modo più intuitivo, mentre la seconda permette di identificare le dimensioni salienti per individuare i cluster nel primo metodo (Seret A., 2012).

Un'altra tecnica di Data mining utilizzata per la segmentazione dei clienti è l'analisi recency, frequency and monetary (RFM) che identifica come clienti di valore quelli che hanno contemporaneamente alti livelli di questi tre elementi, distinguendoli dagli altri. In particolare la recency fa riferimento all'intervallo di tempo intercorrente tra l'ultimo evento, come una transazione, e quello attuale; la frequency rappresenta il numero di eventi di tutta la storia del consumatore e la monetary il valore totale del pagamento (Hu Y., 2012).

I Big Data consentono perciò di guadagnare insight per profilare clienti e prospect efficacemente, permettendo quindi di rivolgersi a target molto più stretti e precisi e di conseguenza di ridurre la dispersione dei contenuti pubblicitari, con ovvi vantaggi anche in termini economici (Casali A., Digital Marketing. Mobile, video, big data e social: internet trasforma la pubblicità, 2015).

La McKinsey propone un metodo che le aziende possono seguire per profilare i clienti sulla base della loro storia Web e sfruttare quindi i Big Data al fine di implementare azioni di Marketing personalizzate. Questo approccio prevede 4 diverse fasi:

- I team di Marketing devono innanzitutto considerare diversi milioni di utenti e la loro storia online, utilizzando cookie o altre forme di tracciamento anonime. Quindi attraverso algoritmi e l'analisi semantica della loro storia Web, possono analizzare i diversi profili basandosi su criteri comportamentali.
- A questo punto i team devono analizzare i comportamenti dei consumatori che hanno acquistato per identificare eventuali correlazioni tra il prodotto e le caratteristiche del profilo dell'utente Web, ricorrendo a tecniche di clustering basate sugli alberi decisionali.
- Segue la costruzione di una campagna digitale personalizzata per il segmento identificato che con maggiore probabilità acquisterà il prodotto.
- Infine gli algoritmi utilizzati devono essere integrati negli strumenti di gestione della digital advertising, in modo che questi diventino parte dei processi quotidiani dell'azienda e vengano attivati solo e automaticamente per gli utenti Web che molto probabilmente verranno convertiti.

Il passo successivo consiste nel ricorso anche ai third party Data, integrati con i first party Data riguardanti i clienti di un certo brand o ricavabili internamente tramite le informazioni dal sito web, app, CRM e mailing list, attraverso le Data Management Platform (DMP). Questo porta allo sviluppo del Data Driven Advertising e alla capacità delle aziende di generare messaggi pubblicitari creativi e personalizzati indirizzati all'utente, utili per dar lui ciò che vuole, dove e quando vuole.

Relativamente all'advertising online e allo sfruttamento dei Big Data, un nuovo fenomeno rilevante è il real time bidding. Le aziende possono utilizzare le grandi moli di dati lasciate dagli utenti che navigano su Internet per scegliere in modo sempre più mirato le campagne display. Una volta che è stato identificato il target di riferimento basta "piazzare la puntata" e

attendere. Il banner del vincitore apparirà esattamente nell'istante in cui l'utente ideale si collega ad un sito del circuito selezionato. Le tariffe vengono determinate attraverso aste condotte su svariati network pubblicitari e l'incontro tra domanda e offerta avviene all'interno di marketplace chiamati Ad Exchange, piattaforme software dove l'advertiser piazza la puntata in real time, che permettono di finalizzare l'acquisto della campagna display in meno di 100 millisecondi e di farla comparire nell'istante in cui l'utente carica la pagina. Questa pratica, per la quale le applicazioni di Data mining rivestono un ruolo molto importante, comporta dei benefici sia per le aziende che per gli utenti. Le prime infatti risparmiano essendo questo un metodo pubblicitario meno costoso, ma allo stesso tempo più mirato, mentre i secondi si vedono apparire offerte a cui potrebbero essere effettivamente interessati anziché essere invasi da pubblicità a random (Tafner C., 2013).

Una nuova tendenza del Digital Marketing è la Visual Listening già discussa nel paragrafo precedente: il listenign delle immagini sta diventando fondamentale dato che al giorno d'oggi sono le foto e i video ad ottenere un livello di engagement molto elevato e non i contenuti testuali, tanto che molte aziende utilizzano Instagram come uno dei più importanti canali di comunicazione.

I vantaggi apportati dai Big Data al Direct Marketing sono, oltre alla personalizzazione del messaggio, la visione a 360° del cliente, l'identificazione dei contenuti, del timing e del canale più appropriato per inviare il messaggio e la possibilità di fare questo in real time. Da ciò deriva un incremento del tasso di conversione, ovvero del numero di visitatori che decidono di cliccare su un certo contenuto casuale o di visitare un sito web come risultato di un'azione guidata, e quindi la massimizzazione del Digital ROI, l'acquisizione di nuovi clienti e la fidelizzazione di quelli che già si rivolgono all'azienda.

### 2.3.2 CUSTOMER MICRO-SEGMENTATION

L'era dei Big Data ha creato la segmentazione 2.0: la molteplicità di nuove tipologie di dati e lo sviluppo di Analytics avanzati permette di avere dettagli granulari e un numero maggiore di informazioni sui consumatori e quindi di generare micro-segmenti molto precisi, costituiti da un piccolo numero di persone. Molti Retailer affermano addirittura di essere impegnati nella personalizzazione e non più nella semplice segmentazione (Manyika J., 2011). I tradizionali segmenti B2C e B2B basati rispettivamente su dati demografici, psicografici e comportamentali e sulle dimensioni delle aziende o sui criteri di acquisto adottati sono ormai superati.

Sfruttando quindi:

- *activity-based Data*, come i clickstream Data dal web, le storie degli acquisti, i dati dei call center, i dati mobile;
- *profili dei social network*, come la storia lavorativa e l'appartenenza a gruppi;
- *sentiment Data*, quindi associazioni a prodotti e aziende (like o follows) e commenti online;
- *dati tradizionali*, come quelli delle ricerche di mercato e quelli transazionali;

è possibile costruire segmenti molto più stretti. Gli uomini di Marketing possono quindi creare offerte, prodotti e servizi personalizzati e su misura per ciascun cluster, con ovvi benefici sui ritorni (Offsey S., 2015). Questi dati possono inoltre essere aggiornati in real time, riuscendo quindi ad identificare i cambiamenti dei clienti e delle loro preferenze.

Emblematico è il caso Amazon, che crea un sito unico per ciascun consumatore: ogni volta che l'utente si registra con il proprio account, Amazon prende nota dei prodotti ricercati o degli acquisti fatti, cosicché la volta successiva che la persona fa accesso, le saranno offerti degli item affini in base alla sua storia passata. (Buckley J., 2015).

La microsegmentazione dei clienti permette quindi di svolgere analisi predittive personalizzate e di ottimizzare le azioni di Marketing: analizzando come queste ultime influenzano i comportamenti di spesa dei diversi microsegmenti, è possibile prevedere i loro diversi livelli di efficacia sui vari minicluster. I benefici che ne derivano sono una maggiore soddisfazione del consumatore e di conseguenza un incremento del suo engagement, una migliore customer experience grazie a contatti personalizzati, la massimizzazione del ritorno sull'investimento di Marketing (MROI) e un aumento del tasso di conversione che porta a ricavi più elevati.

### 2.3.3 PRICE OPTIMIZATION

Le aziende possono sfruttare la crescente granularità dei dati sul pricing e sulle vendite e i potenti Analytics per ottimizzare i prezzi. L'ammontare di informazioni a loro disposizione è enorme, dalle serie storiche della domanda, ai dati relativi alle scorte, a quelli riguardanti i competitor, fino al livello delle vendite attuali. Questa base di dati è in continuo aumento considerata l'esplosione di nuovi canali di vendita online dove i consumatori possono confrontare i prezzi, pratica che ha incrementato la competizione e la frequenza nelle variazioni dei prezzi stessi e considerate le tecnologie digitali e i social media attraverso i quali è possibile tracciare il comportamento di milioni di utenti (Lühr P., 2013). Da queste ingenti quantità di dati, attraverso opportuni tools, i pricing manager sono in grado di estrarre insight per definire

il prezzo ottimale che un consumatore è disposto a pagare per ciascun prodotto quasi in real-time, basandosi sulle sue caratteristiche.

La price optimization può considerare per esempio l'elasticità della domanda al prezzo, con specifici modelli che analizzano i dati delle vendite storiche per ricavare insight sul pricing di ciascuna unità, che possono poi essere utilizzati per fare promozioni o per ridurre i prezzi, valutando i costi conseguenti. Un Retailer del settore alimentare può per esempio considerare l'elasticità al prezzo di consumatori appartenenti a diverse categorie e scoprire come quelli di campagna diano una maggiore priorità a prodotti quali il burro e il riso e siano quindi meno elastici al prezzo per questi rispetto ai consumatori che vivono in città, per i quali sono più importanti alimenti come i cereali e i dolci (Manyika J., 2011).

I benefici che le aziende riescono a conseguire in questo modo sono un aumento dei ricavi, dei margini e della quota di mercato. Tuttavia è necessario seguire alcuni step, descritti di seguito, per riuscire a sfruttare i Big Data in quest'area e raggiungere quindi i vantaggi citati (Baker W., 2014).

- ✓ *Identificare le opportunità più promettenti*, che comprendono determinare quanto il consumatore vuole pagare esattamente per un dato prodotto attraverso la customer segmentation e le promozioni personalizzate.
- ✓ *Ascoltare i dati*: le organizzazioni devono saper far leva sui dati e utilizzare adeguati Analytics per identificare elementi che spesso vengono trascurati, come la situazione economica del Paese e le preferenze dei clienti, e determinare quindi i fattori guida dei prezzi per ciascun cliente e prodotto.
- ✓ *Automatizzare le analisi*: tools automatici sono in grado di individuare segmenti ristretti, determinare cosa assume valore per loro e combinare queste informazioni con i dati transazionali storici. Così facendo, le aziende possono prendere decisioni relative ai prezzi per ciascun prodotto e per ciascun segmento di clienti quasi in real-time, agendo tempestivamente. Automatizzare permette inoltre di replicare le analisi e di effettuare eventuali modifiche molto più velocemente.
- ✓ *Costruire skill e fiducia*: l'implementazione di questo progetto rappresenta non solo una sfida a livello operativo, ma anche una di comunicazione. Le organizzazioni devono innanzitutto convincere i venditori dell'efficacia di questi strumenti nella determinazione dei prezzi e della validità dei suggerimenti offerti. In secondo luogo devono dare una motivazione ai prezzi fissati. Infine i venditori devono essere formati perché questi abbiano fiducia e sappiano dare delle giustificazioni convincenti ai clienti.

- ✓ *Gestire le performance in modo attivo*: le aziende devono fare in modo che il front line abbia visibilità sulla profittabilità dei clienti e che la funzione di Marketing e Vendite abbia le competenze adeguate per cogliere le opportunità offerte dai Big Data in questo ambito e quindi per aggiustare i prezzi, sulla base dei risultati delle analisi.

### 2.3.4 LOCATION-BASED MARKETING

Il Location-based Marketing si basa sull'adozione crescente di smartphone e di altri device mobile che generano i personal location Data, i quali permettono di conoscere dove si trovano le persone in real-time. La figura mostra l'aumento considerevole dei dispositivi utilizzati dal 2013 (dato effettivo) al 2018 (dato previsto) e di conseguenza il notevole incremento del traffico globale di dati mobile (O'Brien B., 2014) .

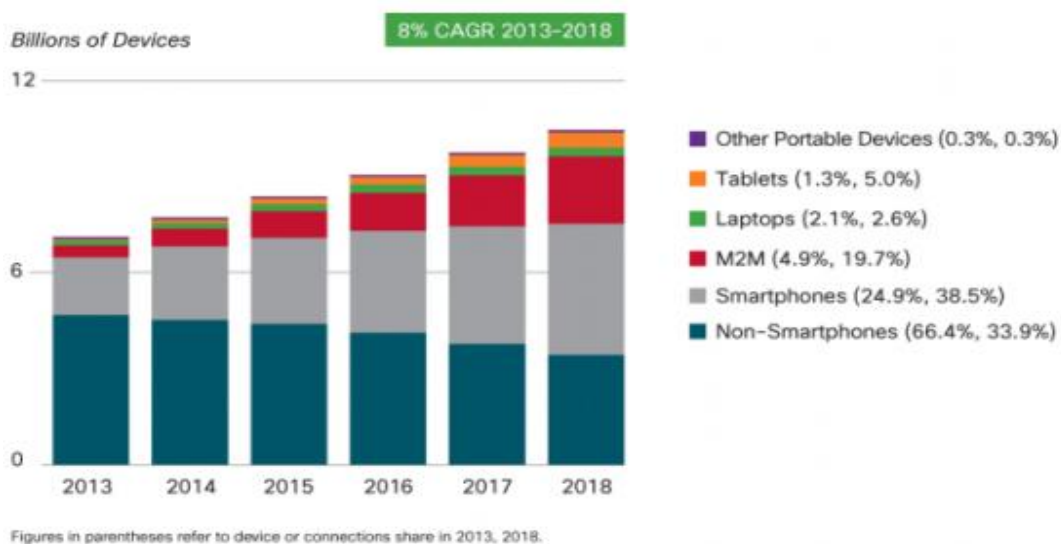


Figura 11: Utilizzo dei dispositivi dal 2013 al 2018 (O'Brien B., 2014)

Quest'ultimo offre infatti alle aziende la possibilità di tracciare geograficamente il comportamento del cliente, favorendo lo sviluppo di una strategia di Marketing che considera le abitudini lavorative e di divertimento e non solo le preferenze dei consumatori.

In particolare sono i GPS negli smartphone e in altri device mobile la fonte principale e in forte espansione di questi dati: essi permettono di identificare la posizione di una persona entro 15 metri utilizzando una costellazione di satelliti orbitanti. Anche sfruttando la capacità di connessione degli smartphone alle reti Wi-Fi è possibile localizzare le persone. Altre due fonti sono i segnali delle torri di triangolazione cellulare e i pagamenti tramite carte di credito e di debito, le quali, attraverso il terminale del punto di vendita, rendono disponibili i dati di

identificazione personale. Infine le tecnologie di prossimità sfruttano le applicazioni come l'RFID e il Near Field Communication (NFC) per comunicare con le persone che si trovano in una determinata area geografica o vicine ad un certo dispositivo. In particolare i tag RFID sono microchip costituiti da una memoria contenente un elevato numero di informazioni e da un'antenna, che attivano un processo di comunicazione e di scambio di dati in lettura e scrittura, identificando l'oggetto in modo univoco. Gli NFC sono una classe dei tag RFID ma hanno una memoria maggiore e contengono molte più informazioni, addirittura un'intera pagina di testo e anche più (Zanotti L., 2015).

Il Location-based Marketing è rivolto a persone che sono in prossimità del negozio o al suo interno. Quello che le aziende fanno solitamente è il Geo-targeted advertising, ovvero effettuare azioni di advertising in tempo reale in base alla localizzazione dei propri clienti. Per esempio i consumatori aderenti possono ricevere sui loro smartphone pubblicità personalizzate relative al loro negozio preferito quando si trovano in prossimità di questo oppure chi è in possesso di uno smartphone e incontra gli amici al bar o al ristorante può ricevere da questi dei coupon per drink e food. Quest'ultima tecnica è il location-based couponing, la quale prevede l'invio di coupon quando i consumatori si trovano in prossimità del negozio. Lo sviluppo di tecnologie che lo rendano possibile è fondamentale in quanto i clienti rispondono positivamente a queste offerte (Alton L., 2015). Diverse aziende americane utilizzano il servizio location-based "push SMS" per spingere i clienti all'interno dei loro negozi. Esso prevede, dopo la scelta dell'area geografica, l'invio di messaggi promozionali e pubblicitari per spronare a visitare un determinato store; ogni utente può ricevere al massimo tre alert a settimana. I risultati sono che la maggior parte delle persone intervistate ha affermato che, quando riceve un SMS significativo, con una maggiore probabilità si recherà nel negozio suggerito (Manyika J., 2011).

Questa forma pubblicitaria è adeguata ai consumatori di oggi: infatti quando questi si recano in un negozio per ricercare un prodotto, utilizzano i loro smartphone per avere più informazioni possibili. Di conseguenza ricevere messaggi rilevanti e personalizzati sui loro device in tempo reale rappresenta una grande opportunità sia per i consumatori che ottengono quello che desiderano quando lo desiderano sia per le aziende che riescono a fidelizzare i clienti.

Per ottenere enormi vantaggi le aziende devono ricorrere al Proximity e Micro-Location Marketing. Vediamo come funziona. Tecnologie emergenti già descritte sopra, quali il Wi-Fi positioning, gli NFC e il GPS degli smartphone permettono di individuare con precisione la posizione del cliente. Una volta che questo è entrato in negozio, gli vengono inviate delle

offerte customizzate mentre cammina con in mano lo smartphone. I negozi devono avere dei dispositivi beacon basati sul bluetooth che inviano informazioni entro brevi distanze, le quali vengono rilevate e processate da app mobile al fine di effettuare azioni, come inviare avvisi e offrire sconti. La Apple ha sviluppato a tal proposito l'applicazione iBeacon. Gli stessi consumatori sono ben disposti a condividere le loro informazioni in cambio di benefici per lo shopping (MarketingLand, 2014).

La push notification tipica delle app mobile rappresenta infatti un importante obiettivo di Location-based Marketing che le aziende intendono raggiungere. Queste vogliono migliorare la capacità del GPS di localizzare le persone in modo da avere più dati possibili sulle loro posizioni per mandare loro avvisi. Oltre ai casi già visti le aziende che organizzano conferenze o eventi importanti potrebbero anche inviare una notifica a tutte le persone interessate nei dintorni del posto in cui si terrà l'evento (Alton L., 2015).

Pertanto lo sfruttamento dei dati di geolocalizzazione a disposizione delle aziende per implementare azioni di Marketing mirate può portare ad un aumento delle vendite, quindi ad un incremento dei profitti e ad un miglioramento della customer experience e perciò alla fidelizzazione della clientela.

Tuttavia relativamente a questo progetto le aziende si trovano a dover affrontare due sfide. La prima è relativa alla privacy: molte aziende utilizzano come vogliono i dati personali riguardanti la localizzazione perché le leggi che devono rispettare per la raccolta, l'aggregazione e la diffusione di queste informazioni non sono chiare. Un framework che descriva quali dati è permesso e quali è proibito usare è necessario per tutti gli stakeholder (Manyika J., 2011). Inoltre c'è un trade-off da considerare in quanto se da un lato i consumatori non vogliono essere osservati e quindi che i loro comportamenti siano tracciati all'interno dei negozi tramite i loro smartphone, dall'altro lato essi desiderano ricevere offerte mobile quando si trovano in prossimità dello store stesso. Risulta perciò utile educare i clienti su come le aziende intendono utilizzare i loro dati (Alton L., 2015).

La seconda sfida è invece tecnologica: settore pubblico e privato devono collaborare per sviluppare device e infrastrutture che siano in grado di generare ulteriori personal location Data.



### 2.3.5 IN-STORE ANALYSIS

L'in-store analysis che prevede l'analisi dei dati relativi al comportamento dei consumatori negli store in real time, è un altro importante progetto Big Data in ambito Marketing.

La posizione e il percorso dei clienti all'interno dello store vengono tracciati attraverso svariate tecnologie: video camere, Wi-Fi e strumenti bluetooth, Wi-fi degli ospiti, sistemi dei punti di vendita, carte di pagamento, trasponder dei carrelli, applicazioni degli smartphone, quali Shopkick e Path Intelligence e tag RFID sulle carte d'acquisto. Ognuno di questi metodi ha diversi livelli di dettaglio: i video per esempio riescono a individuare al meglio il flusso del traffico, ma non il comportamento del singolo individuo, mentre gli RFID sono accurati ma non permettono di identificare i movimenti dell'acquirente in dettaglio, perché per esempio può capitare che un tag attaccato su una carta venga lasciato nel corridoio, mentre il consumatore prosegue nel negozio (Manyika J., 2011).

Tools e Analytics, quali web dashboard, app mobile, real-time alert, strumenti di Data mining, vengono utilizzati per organizzare, analizzare e visualizzare questo grande ammontare di dati, identificare trend e confrontare le prestazioni dei diversi periodi. Così facendo vengono estratti insight relativi ai comportamenti dei consumatori all'interno dello store, con l'obiettivo ultimo di migliorare la customer experience (RetailNext, 2015). In particolare gli insight ottenuti sono relativi a:

- chi è il consumatore: conoscere sesso, età, se è la prima volta che entra nel negozio, se ritorna spesso, da dove viene e quali sono i suoi interessi
- quanti clienti entrano nel negozio: identificare quante persone sono state catturate rispetto a tutte quelle passate davanti al negozio, quanti sono i clienti rispetto ai dipendenti e quanti sono i consumatori entrati un'unica volta
- come si comportano gli shopper all'interno dello store: comprendere come interagiscono con i prodotti, in quali reparti e corridoi si recano e si soffermano maggiormente, in quale ordine visitano le diverse aree, da cosa vengono attratti. In questo modo è possibile determinare le ore in cui ci sono picchi di traffico che generano le principali opportunità di vendita, identificare le vetrine di maggior successo, determinare se i consumatori trascorrono un tempo adeguato nelle diverse zone, capire se le aree con maggior traffico sono quelle che offrono le prestazioni migliori e se le vetrine e i programmi di Marketing stanno generando grandi quote di consumatori, scoprire quanti lasciano il negozio precocemente dopo essere entrati al

fine di ridurre il tasso di abbandono e il tempo necessario per trasformare un prospect in cliente (ShopperTrack, 2015).

Le aziende si servono di questi insight per migliorare efficacemente l'organizzazione, ovvero per ottimizzare il layout dello store, le sue caratteristiche, il posizionamento sugli scaffali e il mix di prodotti offerti per trasformare i clienti una tantum in clienti abituali, incrementare la frequenza delle loro visite e delle loro spese migliorando la store experience, aumentare la dimensione media della transazione e attirare un numero sempre maggiore di consumatori. Le organizzazioni mettono in atto aggiustamenti in tempo reale per ottimizzare l'intero processo d'acquisto.

### 2.3.6 CROSS-SELLING/UP-SELLING

I Big Data offrono grandi opportunità per aumentare la dimensione media dell'acquisto di un consumatore sia mettendo a disposizione prodotti o servizi collegati con la scelta d'acquisto iniziale sia offrendo qualcosa di maggior valore rispetto a questa, ovvero per migliorare le azioni di Cross-Selling e di up-Selling. Dati quali le caratteristiche demografiche dei clienti, la posizione real-time, le preferenze, la storia degli acquisti passati vengono utilizzati a tal fine. Gli algoritmi sviluppati si basano su questi dati per prevedere il comportamento dei consumatori in vari scenari di vendita, estrarre insight per capire molto prima cosa vogliono e determinare quindi il miglior approccio di Cross-Selling e di up-Selling (Baumgartner T., 2011)

I benefici che le aziende traggono sono un aumento delle vendite e quindi dei profitti e la fidelizzazione dei clienti.

Caso esemplare è quello di Amazon che raccoglie i dati da tutti gli utenti, riconoscendo i trend nelle persone che fanno acquisti simili attraverso tools di Analytics, in modo da cogliere potenziali opportunità: se per esempio l'80% degli user che ha comprato scarpe da corsa ha acquistato anche un braccialetto fitness, al prossimo acquisto di scarpe verranno offerti loro anche i braccialetti (Buckley J., 2015). Amazon in base a ciascun prodotto o servizio visitato dall'utente sul sito suggerisce "potresti anche volere" ed è proprio in questo modo che riesce ad incrementare significativamente le vendite.

### 2.3.7 ALTRI PROGETTI

I Retailer online utilizzano i dati di interazione dell'utente con la pagina web, quali lo scrolling e i click per *ottimizzare il design del sito*: è quello che ha fatto eBay con l'obiettivo di determinare

il layout ottimale e altre caratteristiche come la navigazione e la dimensione delle foto. Le aziende inoltre fanno leva sui Big Data, in particolare su quelli transazionali, quelli relativi alle ricerche di mercato e i commenti dei consumatori sui vari social media, per *valutare l'efficacia di nuovi prodotti, servizi o modelli di business* e agire tempestivamente attraverso delle correzioni in real time. Infine le aziende sfruttano i dati storici delle transazioni da svariate fonti e device per *migliorare la market-basket analysis*, individuando i gruppi di prodotti maggiormente profittevoli per poi costruire promozioni, scoprire i driver degli acquisti, ottimizzare il pricing al fine di incrementare la profittabilità e identificare azioni di up-sell e Cross-sell.

### 2.3.8 SINTESI DEI BENEFICI

Dopo aver approfondito i vari progetti, riportiamo sinteticamente nel grafico i benefici apportati dal ricorso ai Big Data e agli Analytics alle organizzazioni nell'ambito del Marketing, emersi in quest'ultima analisi.

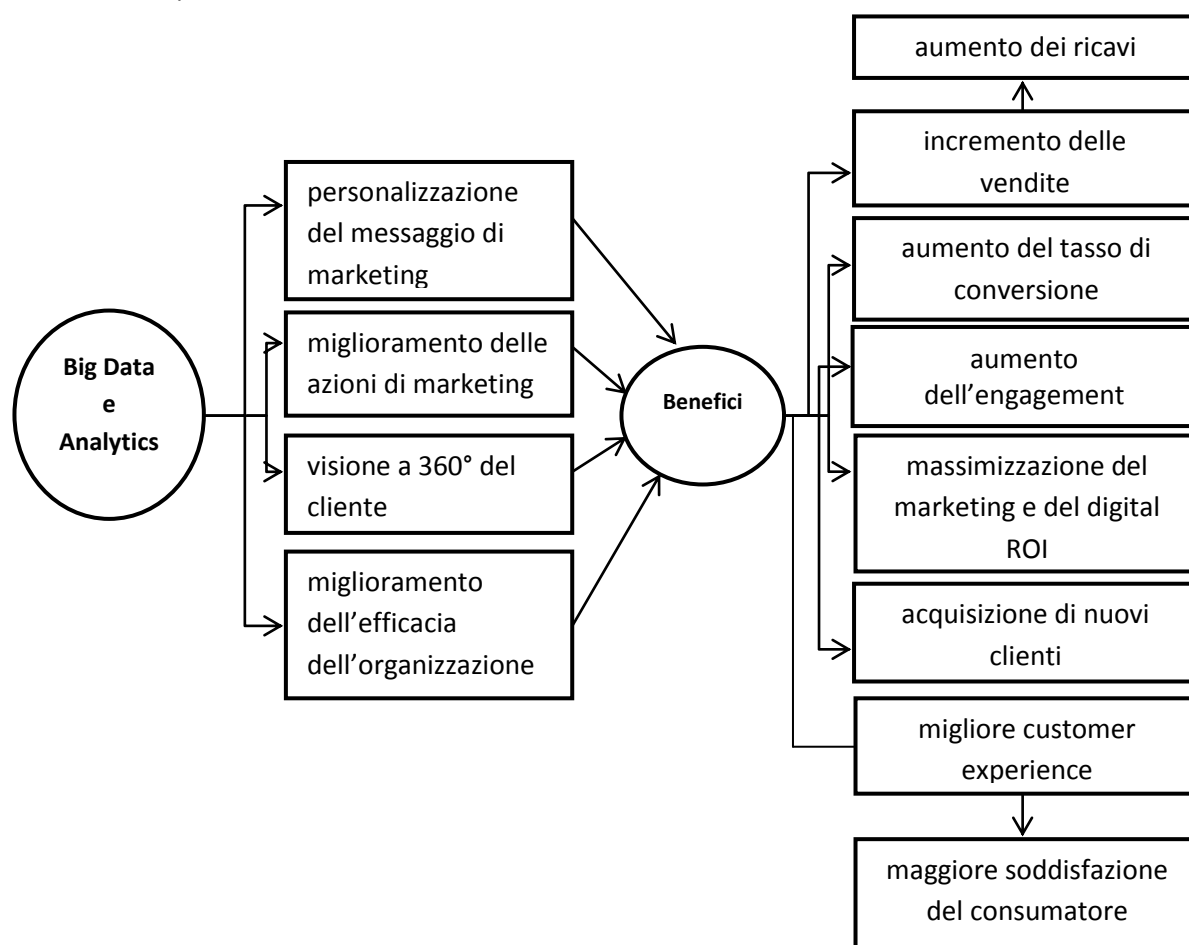


Grafico 6: Sintesi dei benefici apportati dai Big Data in ambito Marketing

## IMPATTO DEI BIG DATA SULLA CUSTOMER EXPERIENCE

Come già sottolineato più volte, una sfida che le aziende devono e vogliono affrontare è lo sfruttamento dei Big Data e degli Analytics per migliorare la *customer experience*, con cui si intende fornire il messaggio o il prodotto giusto alla persona giusto nel luogo e nel momento giusto. Tuttavia si tratta di un compito difficile da raggiungere considerata la nuova complessità del *customer decision journey*: la multicanalità e l'ingente mole di dati generati ad ogni touchpoint rendono ardua l'estrazione di insight relativi al comportamento del consumatore (Fanderl H., 2014). Questi ultimi hanno a disposizione tecnologie che permettono loro di valutare più attivamente prodotti e servizi anche dopo l'acquisto, aggiungendo e rimuovendo continuamente alternative: ciò spinge i brand ad agire in modo attivo per dar forma ai decision journey e non solo a reagire passivamente. Le aziende che adottano questo comportamento riescono a portare valore al brand ma soprattutto al consumatore, garantendo a lui una customer experience migliore. Le fasi di considerazione e di valutazione, che caratterizzano l'approccio classico, vengono ridotte o addirittura eliminate, spingendo il consumer verso la fase di fiducia della relazione, come rappresentato in figura (Edelman D., 2015).

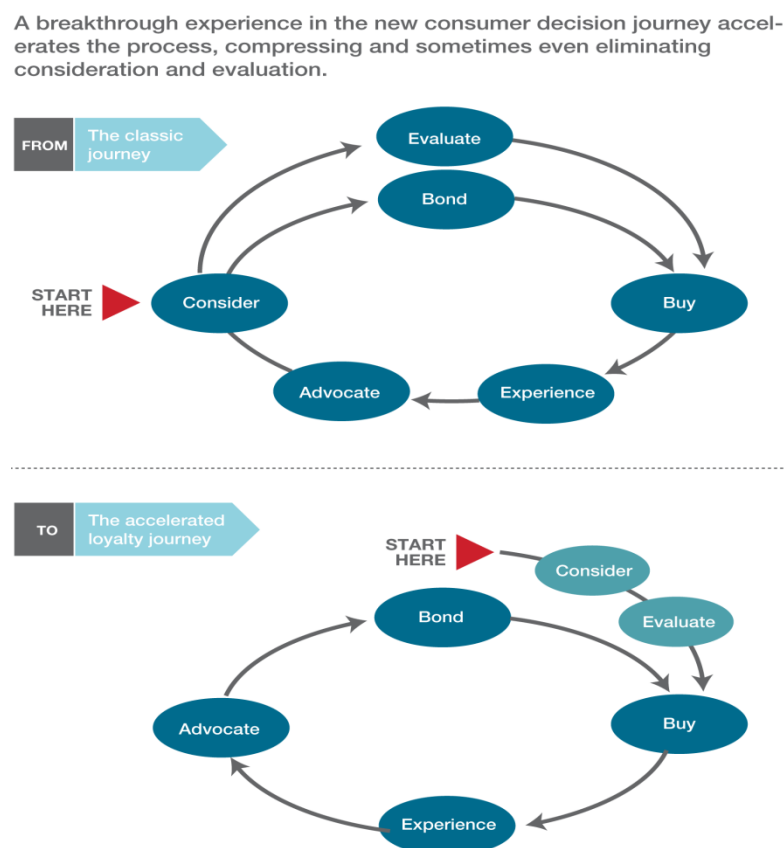


Figura 12: Il classico customer journey e il viaggio accelerato di fedeltà dei clienti (Edelman D., 2015)

La sfida per le aziende consiste nel far leva sui Big Data nel corso dell'intero viaggio decisionale per estrarre insight da integrare nei programmi di Marketing e anticipare i movimenti dei consumatori durante tutto il percorso con l'obiettivo di migliorare la loro esperienza. Le aziende riescono così a trattenere i clienti e questo è fondamentale, considerando che attrarne uno nuovo ha un costo sette volte superiore rispetto a mantenerne uno. L'automazione inoltre viene vista come la capacità critica per fornire una customer experience consistente e personalizzata. Per raggiungere questo obiettivo le imprese devono considerare i tre seguenti elementi (Fanderl H., 2014):

- Focalizzarsi sui top journey: le aziende non devono pretendere di analizzare tutti i bytes di ciascun consumatore, ma dato che i consumer decision journey che contano di più per i consumatori sono 4/5, possono concentrarsi solo sui dati riguardanti questi.
- Non cercare la perfezione: le organizzazioni ritengono di non avere sufficienti dati oppure che quelli che possiedono siano pessimi. In realtà la maggior parte di queste ha i tutti i dati necessari, la difficoltà sta nel metterli insieme.
- Focalizzarsi sui journey Analytics: sono gli Analytics che permettono di identificare cause ed effetti per poi fare predizioni e quindi le aziende devono far leva più su questi che sui reporting, i quali rappresentano solo quello che è successo.

Abbiamo già descritto nel dettaglio come le aziende cercano di sfruttare i Big Data nei diversi ambiti relativi al Marketing per far vivere la migliore esperienza possibile ai consumer. Inoltre nuove innovazioni digitali basate sui dati vengono sviluppate all'interno dei negozi stessi per andare in questa direzione: tra queste riportiamo sistemi di indoor positioning, sistemi di self scanning e carrelli intelligenti, sistemi di cassa evoluti o Mobile POS, sistemi per l'accettazione di pagamenti innovativi e sistemi di loyalty. Esempi sono i programmi fedeltà multipartner forniti da PayBack in tutta Europa che permettono di raccogliere i punti in un'unica carta e di accedere a coupon personalizzati oppure i carrelli intelligenti di Stop&Shop negli USA aventi a bordo sistemi di self scanning che suggeriscono promozioni mirate (Osservatorio Big Data Analytics & Business Intelligence, maggio 2015).

## 2.5 SURVEY ANALYSIS

Riportiamo ora i risultati di alcune survey relative all'utilizzo dei Big Data e degli Analytics in ambito Marketing a conferma di quanto trattato finora in questo secondo capitolo.

Innanzitutto emerge il loro ruolo centrale, tanto che degli oltre 1000 professionisti intervistati da Jon Cifuentes nel 2015 per il report “The State of MarketingAnalytics: Insights in the age of the customer” il 67% ritiene molto importanti i Big Data, il 55% gli audience insight e il 47% i social Analytics. La survey condotta da TeraData e rivolta a 1506 dirigenti di Marketing e di Comunicazione dei principali settori di tutto il mondo, mostra che per il 78% il Data-driven Marketing o è integrato nel sistema aziendale o è strategico e interagisce con il business, con lo sviluppo dei prodotti, con la customer experience e si prevede un aumento di questa percentuale nei prossimi anni. Lo stato sempre più avanzato di questo approccio rivela il potenziale nascosto nei dati, tanto che per l’87% essi sono un asset sottoutilizzato dalle organizzazioni.

I benefici individuati dai vari questionari sono in linea con quelli riportati in precedenza. Decisioni più accurate (67%) e più veloci (59%), migliori risultati di business (57%) e maggiore efficienza (56%) sono i vantaggi più generali che emergono dalla survey di TeraData mentre l’acquisizione e il mantenimento di nuovi clienti (53%), la dimostrazione dell’efficacia dei risultati di Marketing attraverso la misurazione del ROI (45%) e la personalizzazione del messaggio e della customer experience (43%) sono i vantaggi e le priorità più specifiche in ambito Marketing. Inoltre a conferma del fatto che le aziende devono adottare un approccio multicanale, l’83% degli intervistati ritiene che il messaggio customizzato debba essere inviato attraverso svariati canali: i più utilizzati sono i siti web (66%), le app mobile (58%), i call center (51%) e i social media (49%). Un’altra survey del 2014, in cui 500 professionisti della Comunità di Technology Marketing su LinkedIn sono stati intervistati, riporta come i principali benefici di business siano il miglioramento delle performance di Marketing (74%) e una maggiore visibilità del cliente (67%), mentre quelli operativi l’estrazione di insight ricavati dalla combinazione di più fonti (51%), una maggiore velocità decisionale (49%) e la riduzione del tempo sprecato nella gestione di fonti dati frammentate (49%).

Relativamente alle sfide che le aziende si trovano a dover affrontare in quest’ambito, i rispondenti alla survey di TeraData hanno indicato per il 46% la sicurezza dei dati, per il 44% l’assenza di una priorità di investimento e per il 36% la mancanza di commitment e di consenso sull’importanza del Data-driven Marketing. L’indagine proposta da LinkedIn mostra ulteriori ostacoli, distinguendo tra quelli che devono affrontare le grandi organizzazioni (con più di 1000 dipendenti) e le piccole (con meno di 1000 dipendenti). Le prime individuano come maggiori vincoli la mancanza di integrazione nel sistema aziendale (54%) e la qualità e l’integrazione dei

dati (44%), mentre le seconde la mancanza di competenze (33%) e del tempo richiesto per raccogliere e analizzare i dati (32%).

A conferma della barriera relativa alla difficoltà di implementazione delle tecnologie complesse che le aziende devono affrontare, la tecnologia di MarketingAnalytics più adottata dalle aziende sono i fogli di calcolo (66% degli indagati da LinkedIn) per la loro flessibilità, disponibilità e facilità d'uso, seguita da dashboard (54%) e dai tools basati sul cloud e su SaaS (33%). Invece le caratteristiche degli Analytics più importanti sono per il 48% dei rispondenti le dashboard, per il 39% il real-time reporting e per il 34% i predictive Analytics.

Infine a testimonianza della centralità assunta dai social, più della metà delle aziende a cui si è rivolta sempre LinkedIn ha dichiarato di analizzare il social media Marketing e le principali metriche calcolate sono il social reach, ovvero il numero di followers, di like, di membri (62%), il traffico dei social (45%) e le persone maggiormente seguite su questi canali (42%). Si tratta quindi di misure più semplici rispetto ai più complicati Share of Voice e Sentiment Ratio, analizzati finora in piccola misura.

## 3 STATO DELL'OFFERTA

Nei due capitoli precedenti abbiamo fatto emergere le ragioni per cui le aziende dovrebbero sfruttare i Big Data. In questo breve capitolo vediamo se le organizzazioni stanno richiedendo queste soluzioni e se l'offerta di mercato è adeguata.

### 3.1 CONTESTO DI RIFERIMENTO

Oggigiorno le aziende stanno maturando un'esigenza crescente di estrarre insight dall'analisi di enormi quantità di dati di tipologie differenti, andando oltre quella di query e reporting, garantita dalla più tradizionale Business Intelligence. Con quest'ultima ci riferiamo all'insieme di modelli matematici e metodologie di analisi più classiche, comprendenti per esempio la Data Visualization e i Basic Analytics in generale, che esplorano i dati per ricavare informazioni e conoscenze utilizzabili nel corso dei processi decisionali (Vercellis C., 2005). La domanda di Advanced Analytics, che sfrutta sofisticati metodi quantitativi quali il predictive data mining, la simulazione e l'ottimizzazione per ricavare insight al fine di migliorare le performance di business, è in aumento e a inizio anno il mercato sistemato era pari a 1 miliardo di dollari. Il tasso di crescita di quest'ultimo sale continuamente e supera quello della classica BI, il cui valore di mercato è però di gran lunga superiore (Gartner, 2015). Possiamo osservare un analogo trend in Italia: nel 2015 il mercato degli Analytics vale complessivamente 790 milioni di euro, di cui l'84% è relativo alla BI e il 16% ai Big Data, ma, mentre per i secondi è prevista una crescita del 34%, per la prima questa si arresta all'11% (Osservatorio Big Data Analytics & Business Intelligence, 2015). A livello mondiale le principali industry che fanno domanda di Advanced Analytics sono il Retail, l'e-commerce, i servizi finanziari e le comunicazione.

Questi nuovi bisogni hanno in realtà portato ad un cambiamento dell'offerta nell'ambito delle BI e Analytics Platforms. I vendor stanno cercando da un lato di incorporare i classici tool di reporting, dashboard e analisi interattive nei processi di business e dall'altro di porre gli Advanced e Prescriptive Analytics all'interno della piattaforma BI. Per questo stanno introducendo le capacità di data-discovery intelligente, di scoperta dei pattern e di integrazione di dati multistrutturati provenienti sia da fonti interne che da fonti esterne al fine di condurre nuovi tipi di analisi come la location, la sentiment e la graph analysis (Gartner, 2015). L'offerta è molto ampia, così come quella degli Advanced Analytics, per la quale scendiamo maggiormente nel dettaglio, visto che sono le soluzioni che consentono di estrarre il maggior valore dai Big Data.



### 3.2 MAGIC QUADRANT FOR ADVANCED ANALYTICS

Per l'analisi dell'offerta ricorriamo al *Magic Quadrant for Advanced Analytics* di Gartner risalente al 2015, che ci consente di vedere come i player si posizionano in questo mercato. Innanzitutto il Magic Quadrant è uno strumento, risultante da una serie di report relativi a ricerche di mercato di Gartner, che fornisce l'analisi qualitativa di un determinato mercato, mettendo in evidenza aspetti quali la sua direzione, la sua maturità e il posizionamento dei vari attori (Wikipedia, 2015). Nello specifico i due criteri presi in considerazione per la valutazione dei vendor sono *l'ability to execute* e la *completeness of vision*. Il primo fa riferimento alla capacità nel rendere la loro vision una realtà di mercato che i clienti vedono come differenziata dalle altre; fornire una positiva customer experience, supporto, qualità del prodotto, disponibilità di competenze e abilitazione degli utenti determina *l'ability to execute*. Il secondo è invece legato alla loro capacità nel comprendere come le forze dei mercati possano essere sfruttate per creare valore ai consumatori e a loro stessi. In base ai valori assunti da questi due parametri, è possibile individuare 4 diversi casi: i *leaders*, se hanno ottenuto un punteggio alto in entrambe le dimensioni e corrispondono a business maturi e solitamente ad aziende di grandi dimensioni, i *challengers*, se ne hanno preso uno alto nell'*ability to execute* e uno basso nella *completeness of vision* e sono per lo più grandi organizzazioni, *visionares*, se hanno ottenuto punteggi opposti rispetto ai *challengers* e sono solitamente piccole aziende e i *niche players*, se hanno ricevuto uno score basso in entrambi i criteri e sono tipicamente nuove aggiunte nel Magic Quadrant.

Dopo questa breve introduzione allo strumento utilizzato per l'analisi, entriamo nel dettaglio del nostro mercato di interesse, ovvero quello degli Advanced Analytics. Relativamente ai due criteri, *l'ability to execute* è alta se i prodotti e i servizi offerti sono eccellenti in termini di qualità, set di caratteristiche e di competenze; un altro aspetto che impatta su questo criterio è l'esperienza vissuta dal cliente, positiva se i clienti sono soddisfatti del prodotto e della sua integrazione con altri prodotti già esistenti e inviano quindi dei feedback positivi. La *completeness of vision* è elevata se i vendor riescono a capire i bisogni e i desideri dei clienti e a dar forma a questi coerentemente alla loro visione, riuscendo a seguire l'evoluzione del mercato e se il prodotto offerto presenta diverse aspetti di innovazione, tra cui integrazione con R, il supporto ai Data scientist e la Data discovery (Gartner, 2015).

Per essere inclusi in questo quadrante magico i vendor devono però rispettare una serie di requisiti:

- offrire le capacità di Advanced Analytics come prodotto stand-alone o come set di prodotti, i quali devono fornire supporto in diversi aree all'interno dell'organizzazione;
- offrire supporto per le seguenti funzionalità: accesso, preparazione ed esplorazione dei dati; alcune tecniche di predictive Analytics come la regressione lineare e logistica, gli alberi decisionali, le reti neurali e le support vector machine; una serie di tecniche di Descriptive Analytics tra cui il k-means clustering e l'analisi della varianza; testing, validazione e sviluppo dei modelli;
- aver raggiunto un fatturato globale di almeno 3 milioni di dollari negli ultimi 12 mesi derivati da licenze software relative agli Advanced Analytics.



Figura 13: Magic Quadrant for Advanced Analytics 2015 (Gartner, 2015)

Il quadrante magico riportato in figura mette in evidenza la frammentazione del mercato, caratterizzato da un maggior numero di vendor appartenenti al quadrante dei Niche Players. Nello specifico l'offerta è così suddivisa:

1. *Leaders*: sono quattro, *SAS, IBM, KNIME e RapidMiner*. Sono i migliori vendor in questo mercato e guidano la crescita e l'innovazione.
2. *Challengers*: *Dell e SAP* sono le uniche che appartengono a questa classe. Si tratta di due aziende appartenenti ad un mercato adiacente a questo che hanno deciso di entrare nel mondo degli Advanced Analytics e di proporre soluzioni che vengono prese in considerazione dai consumatori. Grazie ad acquisizioni strategiche i due vendor sono riusciti a creare un set di strumenti di grande qualità, al pari dei leader, tuttavia quello che manca rispetto a questi ultimi è la vision chiara di un'offerta integrata. Queste possono diventare due Leaders se riescono ad influenzare il mercato.
3. *Visionares*: questo quadrante è occupato da *Alpine Data Labs, Alteryx e Microsoft*. Questi offrono soluzioni che rispecchiano le tendenze che caratterizzano e caratterizzeranno il mercato, portando vantaggio competitivo ai clienti. Se questi Visionares riuscissero a migliorare la loro Completeness of Vision potrebbero spostarsi nel quadrante in alto.
4. *Niche players*: in questo quadrante troviamo la maggioranza dei vendor, *Angoss, FICO, Predixion, Prognosis, Revolution Analytics, Salford Systems e Tibco Software*. Questi vendor non hanno né la visione integrata dell'offerta dei Leader e dei Visionares né l'ampia portata dei challengers. Alcuni di loro sono dei cosiddetti "Visionaries in attesa" in quanto hanno una vision ma stanno cercando di renderla più convincente, mentre altri possono essere considerati dei "Challengers in attesa" in quanto i loro prodotti non sono ancora così forti ma lo possono diventare continuando a lavorare su di essi.

Per avere una visione più chiara delle soluzioni proposte, riportiamo alcuni esempi di offerta, una per ciascuna categoria che abbiamo descritto.

## SAS

SAS, azienda leader situata nel North Carolina con più di 40000 clienti, è la scelta del maggior numero di organizzazioni che richiedono Advanced Analytics: IDC gli assegna una quota di mercato pari 35,4%, pari al doppio del competitor più vicino. La gamma di applicazioni offerte è ampia: soluzioni di *Data Mining* per la costruzione di modelli predittivi e descrittivi attraverso un'interfaccia grafico che semplifica tutte le fasi che vanno dalla preparazione dei dati fino alla scoperta del modello che fornisce i risultati migliori; *Text Analytics* per trarre vantaggio dall'enorme massa di dati non strutturati quali contenuti testuali provenienti dai Social Media, registri di call center e mail, soluzioni di *Forecasting & Optimization* per capire i trend passati e pianificare il futuro e soluzioni di *Statistical Analysis* per prendere decisioni basate su dati

oggettivi (SAS, 2015). Oltre alla vasta offerta, altri punti di forza sono l'elevata qualità dei prodotti e l'alta scalabilità, stabilità e affidabilità della piattaforma, che contribuiscono alla fidelizzazione della clientela (Gartner, 2015).

### *Dell*

Dell è uno dei due Challengers in questo mercato. Il vendor offre un'ampia gamma di funzionalità: il suo prodotto Statistica permette di risolvere problemi complessi attraverso una piattaforma di Advanced Analytics per migliorare le decisioni di business, di individuare il sentiment dai tweet, dai blog, dai sistemi di CRM e integrarli con i dati demografici per capire meglio il mercato ed identificare le opportunità. Lo sviluppo di questa soluzione è stato possibile grazie all'acquisizione di un'azienda pioniere in questo mercato, StafSoft. Questa piattaforma permette inoltre di integrare fonti di dati strutturati e non e di visualizzare i risultati, individuando i trend (Dell, 2015). I punti di forza della Dell sono le significative risorse finanziarie a disposizione, l'offerta che, come abbiamo già detto, è ampia e il buon livello di customer satisfaction garantito. Tuttavia ci sono anche alcuni punti di debolezza da considerare relativi all'usabilità della soluzione e alla mancanza dell'opportunità di interagire con gli altri clienti (Gartner, 2015).

### *Alpine Data Labs*

Alpine Data Labs occupa il quadrante dei Visioneres. Con sede a San Francisco, offre una piattaforma analitica basata su modelli predittivi, che connette facilmente tutte le fonti di dati, da Hadoop ai Datawarehouse tradizionali (Alpine Data Labs, 2015). I vantaggi sono quindi rappresentati dal fatto che l'azienda fornisce un unico ambiente per gli analisti dei dati e gli utilizzatori finali per la condivisione di idee e feedback e dalla scalabilità della soluzione. Tuttavia gioca a suo sfavore la mancanza di profondità delle funzionalità offerte, in quanto alcune tecniche non sono sviluppate a sufficienza e la minor visibilità sul mercato, date le sue piccole dimensioni rispetto ai grandi vendor (Gartner, 2015). Questi sono i motivi per cui l'azienda si trova nella posizione dei Visioneres nella matrice.

### *Predixion*

Predixion è uno dei tanti Niche Players. L'azienda offre una piattaforma di predictive Analytics basata sul Cloud (è stata la prima azienda ad introdurla) che permette di prendere decisioni migliori al fine di raggiungere ottimi risultati. Questa soluzione offre supporto dalla preparazione dei dati allo sviluppo del modello ed è in grado di eseguire gli Analytics sulla

piattaforma di BI, sui portali, sui device mobile e sulle applicazioni di business (Prexion, 2015). I prodotti di Predixion ricevono feedback positivi dagli end-user e si contraddistinguono per la facilità di sviluppo in ambienti diversi, come appena spiegato. Gli svantaggi sono rappresentati dalla scarsa visibilità dell'azienda sul mercato, dal complesso sviluppo iniziale e dall'offerta specifica focalizzata verso poche industry (Gartner, 2015).

## 4 METODOLOGIA DI RICERCA

### 4.1 INTRODUZIONE

In questo capitolo illustriamo il percorso metodologico che ha condotto alla realizzazione di questo lavoro di tesina. Nella prima parte accenniamo all'Osservatorio Big Data Analytics & Business Intelligence, all'interno del quale si collocano le analisi e gli studi effettuati, in seguito esponiamo la metodologia utilizzata, dallo studio della letteratura all'analisi dei risultati estratti dai dati rilevati secondo diverse modalità.

### 4.2 OSSERVATORIO BIG DATA ANALYTICS & BUSINESS INTELLIGENCE

Questo lavoro di tesina è stato svolto nel contesto della Ricerca 2015 condotta dall'Osservatorio Big Data Analytics & Business Intelligence della School of Management del Politecnico di Milano, il cui obiettivo è mettere in evidenza il valore strategico assunto dalle metodologie di Business Intelligence e dai Big Data Analytics per le imprese e la Pubblica Amministrazione e i potenziali benefici che ne derivano, consistenti in un aumento della competitività, della redditività, della tempestività e dell'efficacia dei processi decisionali. L'Osservatorio intende diventare un punto di riferimento per i CIO e i manager di linea per la conoscenza e la selezione di queste principali soluzioni affinché riescano ad ottenere innumerevoli vantaggi.

Nato nel 2008, l'Osservatorio si è sin da subito concentrato su moltissimi settori, tra cui Fashion, Manufacturing, Retail, Media, Telco, Utilities, Sanità, Banking & Finance e GDO & Logistic, focalizzandosi maggiormente negli ultimi anni sul Marketing e la Customer Experience. Più nello specifico gli obiettivi della Ricerca sono la stima del mercato e lo stato di diffusione delle soluzioni, l'identificazione del loro grado di penetrazione nelle diverse linee di business, la definizione di un "Big Data Journey" che determini i percorsi e i passi per l'adozione, l'analisi dei benefici, delle barriere, di nuovi ruoli organizzativi e dell'offerta delle startup e l'indagine delle applicazioni in specifici settori al fine di individuare i casi di successo.

Anche quest'anno l'Osservatorio ha condotto l'iniziativa Big Data Innovation Award, finalizzata a sostenere la cultura dell'innovazione in ambito Big Data, dove vengono premiate le aziende che hanno introdotto i migliori e più innovativi sistemi di Big Data Analytics & Business Intelligence, differenziandosi dalle altre organizzazioni.

I risultati delle ricerche e delle survey estese ad aziende italiane appartenenti a tutti i settori merceologici vengono riportati in report cartacei ed elettronici annuali e attraverso diversi convegni ai quali partecipano i principali protagonisti del mondo delle imprese e i rappresentanti dell'offerta di sistemi di Big Data Analytics & Business Intelligence.

### 4.3 PERCORSO DI RICERCA

Nel seguente grafico riportiamo gli step seguiti per la stesura del lavoro.

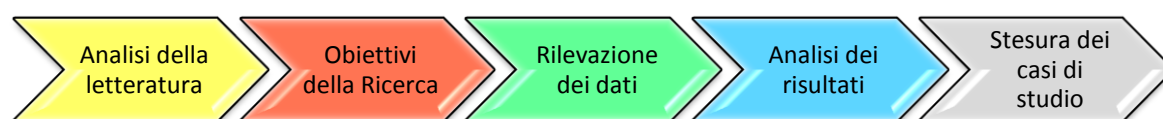


Figura 14: Passi determinanti il percorso di Ricerca

Il primo passo è stato analizzare la letteratura, dove abbiamo definito il contesto per le successive analisi, utile per la validazione dei risultati delle survey e delle interviste; di seguito abbiamo individuato gli obiettivi della ricerca, rilevato i dati, che sono stati quindi analizzati. L'ultima fase ha previsto invece la stesura di una serie di casi di successo dopo aver sostenuto interviste telefoniche.

#### 4.3.1 ANALISI DELLA LETTERATURA

Il lavoro di tesina ha preso le mosse dall'analisi della letteratura di riferimento sul tema Big Data, del quale abbiamo approfondito svariati aspetti. In particolare abbiamo trattato le principali caratteristiche, i benefici che possono essere tratti dal loro utilizzo, le barriere che ostacolano il ricorso a questi, le tecniche e le tecnologie di Big Data, i trend in atto e la governance dei sistemi. A tal fine abbiamo analizzato le principali pubblicazioni a livello internazionale, di cui le più importanti sono state i report della *McKinsey* e gli articoli estratti da innumerevoli fonti, tra cui il *McKinsey Quarterly*, l'*Harvard Business Review* e il *Browse Journal & Magazines*. Inoltre abbiamo utilizzato i dati delle presentazioni mostrate negli workshop tematici relativi ai Big Data Analytics & Business Intelligence, che si sono tenuti nel corso dell'anno. Abbiamo quindi illustrato più in dettaglio l'effetto dei Big Data nei settori della Sanità, della Pubblica Amministrazione, del Manufacturing e del Retail, ricavando le informazioni da diversi report e dal materiale online riguardante il loro utilizzo nello specifico ambito analizzato.

Nel secondo capitolo ci siamo concentrati sull'impatto dei Big Data sul Marketing e sulla Customer Experience, trattando nello specifico i Marketing Analytics e i progetti in quest'area, con un focus particolare sui Social Analytics. Per questa parte abbiamo esplorato svariate fonti, tra cui articoli con un taglio maggiore sul Marketing tratti da riviste e periodici come il *McKinsey Quarterly*, l'*International Journal of Research in Marketing* e il *Decision Support Systems*. Alcuni aspetti sono stati dedotti dallo studio di documenti e informazioni trovate su siti specializzati sul tema, di cui i principali sono quello di *BIT BANG*, società leader nella consulenza Web Analytics e Customer Experience Management, *Digital4*, primo progetto culturale italiano per gli Executive di Business focalizzato sull'impiego strategico delle tecnologie digitali della società editoriale ICT&Strategy e quello di *IBM*.

Per una prima conferma dei risultati e dei temi discussi nei primi due capitoli abbiamo analizzato i dati di alcune survey tratte dal web, rivolte a CIO e CMO di imprese di tutto il mondo e condotte da *Gartner*, *IDG Enterprise*, *Wikibon*, *Forbes*, *TeraData* e *Linkedin*.

Nel terzo capitolo ci siamo concentrati sull'analisi dell'offerta attuale in ambito *Advanced Analytics Platforms*, che rappresentano gli strumenti che permettono di estrarre maggiormente valore dai Big Data. A tal fine abbiamo preso come riferimento il *Magic Quadrant for Advanced Analytics* di *Gartner* aggiornato all'anno corrente e approfondito l'offerta di alcuni vendor individuati utilizzando le informazioni trovate sui rispettivi siti.

#### 4.3.2 OBIETTIVI DELLA RICERCA

A partire dall'analisi della letteratura abbiamo quindi identificato gli obiettivi del lavoro di tesina, che sintetizziamo di seguito:

- fornire una definizione di Big Data, considerando i diversi aspetti caratterizzanti e il contesto in cui si inseriscono;
- individuare il livello di maturità delle medie e grandi organizzazioni italiane ai Big Data Analytics;
- indagare l'approccio ai Big Data Analytics adottato dalle diverse funzioni aziendali, identificando il loro stato di maturità in quest'ambito;
- analizzare il ruolo delle start-up sia a livello internazionale che a livello nazionale nell'ambito dei Big Data Analytics individuare il supporto fornito;
- esaminare i progetti Big Data all'interno delle funzioni aziendali più mature;
- identificare i casi di successo.



Il secondo, terzo, quarto e quinto obiettivo rappresentano le domande di ricerca da cui siamo partiti per questo lavoro di tesina.

#### 4.3.3 RILEVAZIONE E ANALISI DEI DATI

Questo lavoro di tesina prende le mosse dalle quattro seguenti domande di ricerca, che abbiamo già citato nel paragrafo precedente:

1. Quale è il livello di maturità delle medie e grandi organizzazioni italiane ai Big Data Analytics?
2. Quale è il livello di maturità delle funzioni aziendali all'interno delle medie e grandi organizzazioni italiane ai Big Data Analytics?
3. Quale è il supporto dell'offerta di mercato?
4. Quali sono gli ambiti progettuali delle funzioni aziendali risultate avere il maggior livello di maturità?

Per rispondere alle domande 1-2-4 ci siamo serviti dei dati che abbiamo rilevato attraverso due survey online che hanno coinvolto rispettivamente 160 C-level di diverse funzioni aziendali e 91 CIO e Responsabili IT di medie e grandi organizzazioni e attraverso interviste ai Responsabili di progetto di iniziative di BDA. Un aspetto critico della prima è rappresentato dal campione non molto consistente per le diverse funzioni aziendali, ad eccezione del Marketing e Vendite. Ci siamo inoltre serviti della survey dell'anno scorso rivolta a 73 Responsabili ed Executive Marketing sempre di medie e grandi organizzazioni per un confronto per l'ultima domanda di ricerca.

Ai fini della nostra indagine e quindi per condurre un'analisi più approfondita delle singole funzioni ci siamo concentrati maggiormente sulla prima, in cui gli argomenti trattati sono stati:

- il pensiero relativo a cosa rappresentano i Big Data;
- la tipologia di dati utilizzati dalle funzioni aziendali e le loro fonti;
- le soluzioni di Big Data Analytics adottate e le motivazioni e i freni alla base della loro implementazione;
- le mansioni e i progetti in cui i Big Data vengono sfruttati nelle diverse funzioni aziendali.

Riportiamo nel seguente grafico i ruoli organizzativi ricoperti dagli intervistati.

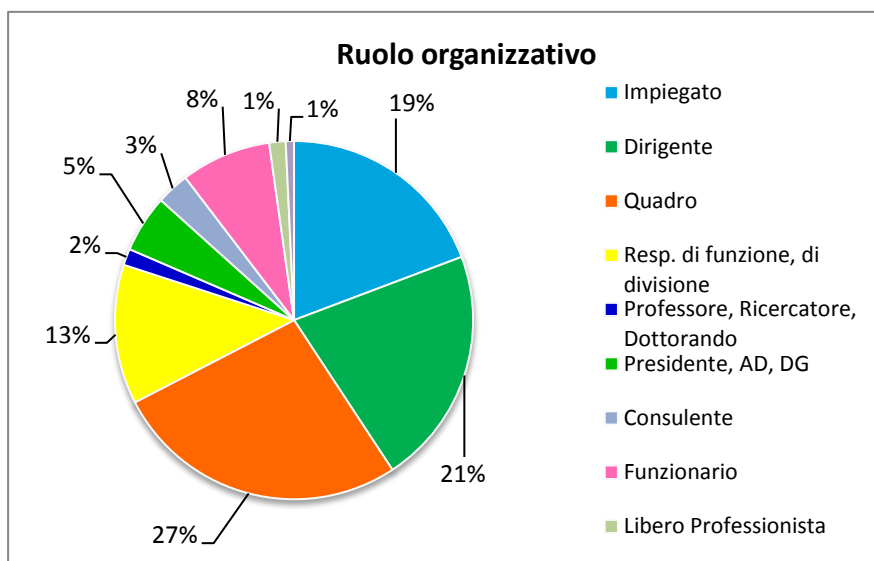


Grafico 7: Scomposizione campione survey C-level per ruolo organizzativo

I temi indagati nel questionario rivolto ai CIO e ai Responsabili dei Sistemi Informativi sono stati invece:

- la strategia dell'organizzazione verso l'adozione dei sistemi di BDA&BI;
- la tipologia di dati analizzati da queste soluzioni;
- la governance dei sistemi;
- le scelte di sourcing e gli ambiti di applicazione.

Per la prima domanda di ricerca ci siamo concentrati sul primo punto.

I questionari hanno rappresentato per noi uno strumento fondamentale, che ha permesso di ottenere grandi quantità di informazioni e di condurre un'indagine quantitativa tramite domande e risposte strutturate. A partire dai risultati tratti dall'analisi effettuata a valle sono state selezionate una serie di iniziative interessanti alle quali se ne sono aggiunte altre rilevate attraverso la ricerca su fonti secondarie o segnalate dai partner dell'Osservatorio. Queste sono state in seguito approfondite attraverso delle interviste telefoniche ai responsabili del progetto che ci hanno dato la possibilità di svolgere un'analisi qualitativa, più flessibile e di raccogliere ulteriori informazioni e dettagli, che per la natura rigida del questionario non potevano essere colti. Prima del loro svolgimento è stata stesa una traccia da seguire, i cui elementi chiave sono le ragioni per cui le aziende hanno iniziato ad affacciarsi al mondo Big Data e l'approfondimento di un'iniziativa implementata sia da un punto di vista organizzativo che da

uno tecnologico, con focus specifico sui benefici e le criticità affrontate. Il giusto mix di questi di survey e call si è rivelato quindi una scelta vincente.

Entriamo ora nello specifico del procedimento seguito per le singole domande di ricerca. Innanzitutto per identificare il livello di maturità delle aziende ai Big Data Analytics (prima domanda di ricerca) abbiamo analizzato la *survey CIO*, considerando in particolare lo stato di adozione dei BDA e l'approccio seguito per la gestione del patrimonio informativo. Abbiamo quindi mappato su una matrice le due dimensioni per una classificazione delle aziende in 4 categorie.

Per determinare il livello di maturità delle funzioni aziendali ai BDA (seconda domanda di ricerca) siamo ricorsi alle *survey rivolte ai responsabili C-Level delle funzioni aziendali* e *CIO* e alle *interviste*, procedendo attraverso i seguenti passi:

1. *Costruzione di un cruscotto di 10 indicatori*: 1. Livello di fruizione attuale; 2. Potenzialità future; 3. Consapevolezza dell'impatto strategico dei Big Data; 4. Utilizzo di dati semi/destrutturati; 5. Utilizzo di dati esterni; 6. Qualità della gestione dei dati; 7. Adozione di tool avanzati di analisi dei dati; 8. Numero di tipologie di soluzioni adottate per l'analisi dei dati; 9. Grado di adeguatezza delle soluzioni.
2. *Attribuzione ad ogni indicatore, per ciascuna funzione aziendale, di un livello di maturità su una scala da 1 a 5* (1 rappresenta il livello più basso e 5 quello più alto), rappresentato da uno specifico colore:
  - ✓ Livello 5 (verde scuro)
  - ✓ Livello 4 (verde chiaro)
  - ✓ Livello 3 (giallo)
  - ✓ Livello 2 (arancione)
  - ✓ Livello 1 (rosso)
3. *Determinazione del livello di maturità di ogni funzione aziendale*, attraverso il calcolo della media dei livelli attribuiti a tutti gli indicatori per ciascuna di esse.

Per individuare gli ambiti progettuali delle funzioni aziendali più mature che abbiamo individuato nel punto appena descritto (quarta domanda di ricerca) abbiamo utilizzato le *survey rivolte ai responsabili C-Level delle funzioni aziendali* e le *interviste*.

Infine per rispondere alla domanda relativa al supporto dell'offerta di mercato alla richiesta di soluzioni di BDA delle funzioni aziendali, abbiamo utilizzato come fonte principale il portale

*Crunchbase*, da cui abbiamo tratto la maggior parte delle aziende di interesse e come fonti secondarie *altri siti*, quali Google. Precisamente abbiamo proceduto nel modo seguente:

1. *Censimento di 498 startup* italiane e internazionali che hanno ottenuto finanziamenti dal 2012. Gli aspetti considerati sono i dati anagrafici (data di nascita, continente e nazione headquarter, descrizione, sito web, scope geografico, contatti, principali investitori, industry di riferimento, Osservatorio di competenza), i finanziamenti complessivi ottenuti e i finanziamenti ottenuti dal 2012. Abbiamo quindi classificato le startup in base alla tipologia e all'origine dei dati analizzati e agli ambiti applicativi offerti. Relativamente a quest'ultima dimensione, che è quella più importante, abbiamo individuato 3 categorie: *Infrastructure*, a sua volta suddivisa in Database, Technologies, Storage e Cross-infrastructure, *Analytcs* per la quale abbiamo identificato le classi di Data Visualization, Search, Location/people/events, Analytcs platform e Social Analytcs e *Application*, suddivise negli ambiti di Marketing&Sales, CRM & Customer Experience, Finance, Supply Chain, Logistica & Operations, IT, Human Resource e Security, per un totale quindi di 16 categorie. Infine ci siamo focalizzati sulle startup italiane, per le quali abbiamo analizzato anche la figura del/dei fondatori, considerando nello specifico il profilo di competenze e l'età.
2. *Identificazione del livello di offerta* per ciascuna funzione aziendale utilizzando i dati delle startup appena citati, attraverso una tabella.
3. *Identificazione del livello di domanda* per ciascuna funzione aziendale utilizzando i dati della *survey CIO*, attraverso una tabella.
4. *Mappatura su una matrice dello stato di offerta e di domanda* di ciascuna funzione aziendale e spiegazione

Brevemente, i risultati emersi sono i seguenti:

- le imprese non hanno ancora raggiunto un alto livello di maturità nell'ambito dei Big Data Analytics e si trovano ancora nella fase iniziale del percorso che porterà ad un pieno sfruttamento di queste soluzioni;
- il Marketing e Vendite è risultato la funzione maggiormente pronta ad adottare i sistemi di BDA e a trarne vantaggio, seguita dall'Amministrazione, Finanza e Controllo. Nel complesso possiamo osservare funzioni come queste che sono già in una buona posizione e altre per le quali c'è ancora molto lavoro da fare;

- l'offerta delle startup censite è ampia e soddisfa le esigenze delle funzioni aziendali. È ancora la funzione di Marketing e Vendite ad avere a disposizione il maggior supporto anche in quest'area;
- i progetti che mirano allo sfruttamento dei Big Data sono già presenti all'interno delle funzioni aziendali più maturi, in cui riscontriamo anche un enorme interesse prospettico per altre iniziative. Per il Marketing queste vanno dal Direct Marketing ai Social Analytics, dal Location Based Marketing fino al Price Optimization e molto altro.

Illustriamo in un grafico, riportato nella pagina seguente il percorso di analisi appena descritto, per fornire una visione più chiara e sintetica

#### 4.3.4 STESURA DEI CASI DI STUDIO

L'ultima parte di questo lavoro di tesina prevede la stesura di 9 casi di studio relativi a progetti di Big DataAnalytics di particolare interesse. La scelta è stata fatta, come già detto, sulla base delle risposte date dalle aziende alle survey proposte e agli interventi dei loro rappresentanti nel corso degli workshop che si sono tenuti durante l'anno.

Questi casi rappresentano uno strumento molto utile perché permettono di scendere più nel dettaglio rispetto alle risposte dei questionari relativamente a quanto è stato implementato in azienda e di avere una visione più concreta dei temi trattati. Nello specifico le fonti utilizzate per l'elaborazione dei casi sono le seguenti:

- interviste telefoniche, che ci hanno permesso di ottenere le informazioni desiderate e di approfondire maggiormente le iniziative;
- survey, attraverso le quali abbiamo raccolto i dati quantitativi;
- interventi nei vari workshop in cui sono stati esposti alcuni dei progetti più rilevanti;
- siti internet delle aziende e altri in cui sono state trovate informazioni relative al caso trattato;
- materiale, in particolare presentazioni in powerpoint, che le aziende hanno messo a disposizione.

Di seguito riportiamo la struttura seguita per la redazione dei casi:

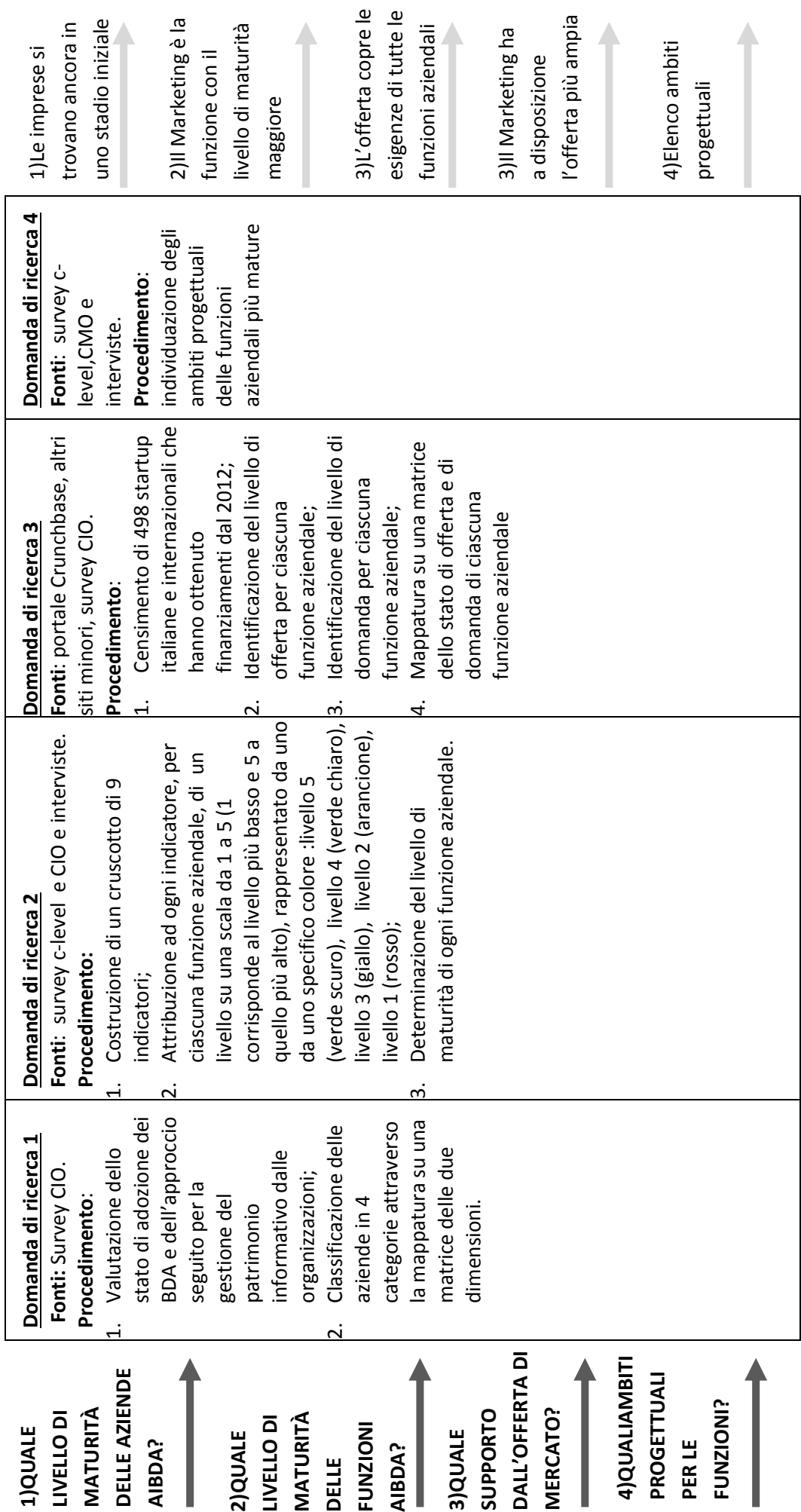


Grafico 8: Schema del percorso di ricerca seguito

- breve descrizione dell'azienda;
- motivazione ed esigenze alla base dell'introduzione di una soluzione di Big Data;
- descrizione dell'iniziativa in termini di fasi del progetto, funzionalità della soluzione, ambiti, governance e fattori alla base del suo successo;
- benefici apportati e previsti dall'introduzione della soluzione;
- criticità affrontate durante l'introduzione dell'iniziativa;
- progetti futuri che l'azienda intende implementare.

#### 4.3.5 WORKSHOP

Infine per approfondire i temi indagati abbiamo fatto riferimento ai risultati ottenuti dai vari workshop che si sono tenuti nel corso dell'anno dall'Osservatorio Big Data Analytics & Business Intelligence. Nello specifico il workshop *"I Big Data come leva strategica nel Retail"* che aveva l'obiettivo di analizzare le opportunità offerte ai retailers dalle enormi quantità di dati prodotti e gestiti ogni giorno, sia quelli transazionali che quelli provenienti da fonti esterne alle organizzazioni; il workshop *"I Big Data nell'online business"*, che intendeva analizzare le opportunità che le nuove tecnologie offrono per raccogliere, gestire e processare grosse quantità di dati in formati diversi, con particolare riferimento alle organizzazioni che operano nell'online business; il workshop *"I Big Data per la customer journey"* che aveva l'obiettivo di indagare come i Big Data possono impattare efficacemente sulle strategie di Marketing ed in particolare sulla Customer Experience. Ulteriori spunti sono stati tratti dal convegno conclusivo tenuto dall'Osservatorio Big Data Analytics & Business Intelligence, dove sono stati presentati i risultati della Ricerca, approfonditi diversi temi tra cui gli scenari di innovazione per le imprese, l'ecosistema startup ed esposte alcune esperienze di successo nelle organizzazioni italiane da parte dei responsabili di progetto.

## 5 ANALISI DEGLI APPROCCI AI BDA DELLE FUNZIONI AZIENDALI E DEL SUPPORTO FORNITO DALL'OFFERTA

In questo capitolo cerchiamo di rispondere alle domande di ricerca da cui è partito il nostro studio, che riportiamo brevemente di seguito:

- Quale è il livello di maturità delle medie e grandi organizzazioni italiane ai Big DataAnalytics?
- Quale è il livello di maturità delle funzioni aziendali all'interno delle medie e grandi organizzazioni italiane ai Big Data Analytics?
- Quale è il supporto dell'offerta di mercato?
- Quali sono gli ambiti progettuali delle funzioni aziendali risultate avere il maggior livello di maturità?

Le principali fonti dati utilizzate sono state un'intervista rivolta a 91 CIO di medie e grandi organizzazioni italiane, una rivolta a 160 C-level sempre di medie e grandi aziende italiane, interviste ad alcuni responsabili di progetti BDA e per l'indagine sulle startup il portale Crunchbase.

### 5.1 LIVELLO DI MATURITÀ DELLE AZIENDE AI BIG DATA ANALYTICS

Dall'analisi della survey rivolta a 91 CIO di medie e grandi organizzazioni italiane e ad una serie di interviste abbiamo cercato di individuare il livello di maturità delle aziende ai Big DataAnalytics.

È vero che i Big Data stanno sempre di più attirando l'attenzione delle organizzazioni, tuttavia per ora si tratta per lo più di speranze e possibili opportunità e non di risultati concreti e maturi. L'analisi del grado di adozione e di pervasività dei sistemi Big Data ne è una conferma: nella maggioranza delle organizzazioni, il 42%, le soluzioni sono ancora in fase di valutazione, nel 27% ci sono dei progetti pilota in alcuni ambiti specifici, nel 16% c'è un utilizzo a regime ma solo in determinate aree e non nella maggior parte e nel 14% le soluzioni sono assenti e non c'è nessun interesse ad introdurle.



L'approccio seguito dalle organizzazioni per la gestione del patrimonio informativo è infatti tradizionale: nel 33% del campione non c'è un piano condiviso e il potenziale di questi sistemi non è ancora stato compreso completamente, mentre nel 27% sono presenti iniziative gestite dalle singole funzioni, che hanno specifiche esigenze e comprendono i vantaggi derivanti dai Big Data. Il fatto che il 60% gestisca in modo classico la conoscenza aziendale conferma dunque lo stadio iniziale in cui si trovano le aziende. Ce ne sono comunque alcune che hanno fatto dei passi avanti, sebbene queste rappresentino una piccola porzione: il 16% ha un piano annuale condiviso dal top Management che crede nelle opportunità offerte da queste soluzioni, il 18% ne ha uno pluriennale con KPI condivisi e continui investimenti previsti e solo il 6% considera i Big Data parte integrante della strategia complessiva dell'azienda.

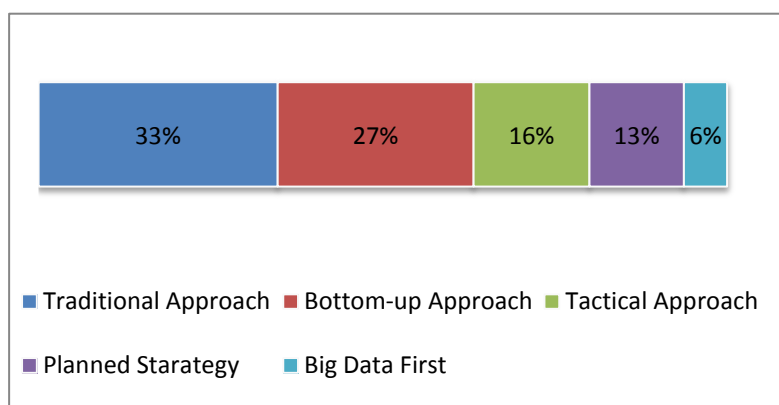


Grafico 9: Approccio delle organizzazioni alla gestione dei Big Data

I dati mettono in evidenza come nel complesso il mondo delle medie e grandi aziende si trovi ancora in uno stadio iniziale in quest'ambito: se da un lato infatti ci sono organizzazioni che stanno cercando di far leva sui Big Data a 360°, consci delle opportunità strategiche che questi offrono, dall'altro ce ne sono alcune maggiormente sospettose, che risultano quindi frenate.

Mappando su una matrice, che riportiamo nella pagina seguente, lo stato di adozione dei Big Data Analytics, ipotizzando come possibili casi "soluzioni assenti" e "soluzioni presenti", e l'approccio seguito, distinguendo tra "tradizionale" (che non prevede un piano condiviso e non coglie il potenziale dei Big Data) ed "orientato allo sfruttamento dei Big Data", possiamo identificare quattro principali tipologie di organizzazioni:

- *pionieri* (27%): riconoscono l'importanza dei Big Data per la strategia aziendale e sono in una fase avanzata di implementazione degli Analytics;

- *reattivi* (16%): sviluppano un piano condiviso per la gestione dei Big Data essendo consapevoli dei vantaggi che questi possono apportare, ma sono in ritardo nell'adozione delle soluzioni;
- *tradizionalisti* (15%): non hanno una mentalità aperta verso le possibilità che i Big Data mettono a disposizione all'intera organizzazione, tuttavia implementano dei sistemi in specifici ambiti;
- *scettici* (40%): non hanno ancora compreso a pieno le opportunità offerte dai Big Data e per questo motivo non hanno molto interesse ad introdurre strumenti di Big DataAnalytics.

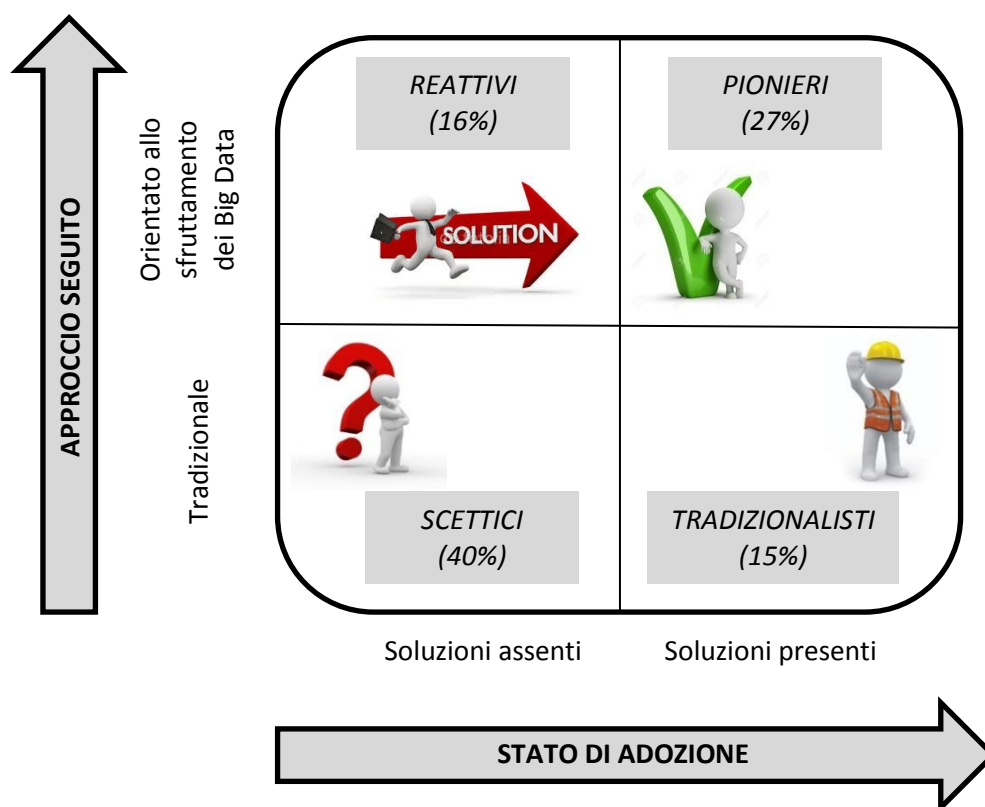


Grafico 10: Matrice stato di adozione dei Big Data Analytics e approccio seguito dalle organizzazioni

Comunque l'interesse in crescita verso questo tipo di soluzioni è testimoniato dal significativo aumento rispetto al 2014 del budget a loro destinato, pari al 34%. Si tratta di un tasso triplo rispetto a quello della Business Intelligence, pari all'11%, che testimonia appunto un'attenzione sempre maggiore rivolta ai Big Data piuttosto che alla tradizionale BI.

## 5.2 LIVELLO DI MATURITÀ DELLE FUNZIONI AZIENDALI AI BIG DATA ANALYTICS

Cerchiamo ora di valutare se le diverse funzioni aziendali sono pronte ad adottare e a trarre i conseguenti vantaggi dalle soluzioni di Big Data Analytics, sfruttando i dati della survey rivolta a 160 C-level di medie e grandi organizzazioni italiane. Le conclusioni tratte derivano dall'integrazione di questa fonte con le informazioni ricavate dalle interviste ai responsabili di diversi progetti Big Data e con ragionamenti personali logici visto che, considerando il campione non significativo per alcune funzioni aziendali, certi risultati potevano essere fuorvianti.

Per raggiungere il nostro obiettivo abbiamo costruito innanzitutto un cruscotto di 9 KPI, che elenchiamo di seguito:

1. *Livello di fruizione attuale*: grado di utilizzo dei sistemi di BDA&BI all'interno della funzione aziendale.
2. *Potenzialità future*: livello di opportunità previsto per il futuro delle soluzioni di BDA&BI per la funzione aziendale.
3. *Consapevolezza dell'impatto strategico dei Big Data*: convinzione che i Big Data rappresentino un fenomeno rivoluzionario o comunque un trend rilevante che influenzerà enormemente il modo di fare business di un'impresa, piuttosto che solo un sostegno per lo svolgimento delle attività della singola funzione. Si tratta quindi di un indicatore che ci permette di verificare se ciascuna funzione ha realmente compreso l'importanza e l'effetto dei Big Data sull'organizzazione.
4. *Utilizzo di dati semi/destrutturati*: livello di sfruttamento di dati "non convenzionali" e quindi o privi di schema o con una struttura irregolare o parziale. Il ricorso a questi è molto importante dato che il 90% dei dati presenti nell'universo digitale è di natura destrutturata.
5. *Utilizzo di dati esterni*: livello di sfruttamento di dati provenienti da fonti esterne come quelli dal Web, molto importante considerando l'enorme mole di dati generata da questo mondo. In particolare con questo KPI vogliamo identificare il livello di analisi di dati provenienti dai Social Media, considerata la loro esplosione e la loro importanza sempre maggiore, motivo per cui le funzioni dovrebbero dedicarsi al loro studio.
6. *Qualità dei dati*: livello di qualità e di affidabilità nell'immagazzinamento e messa a disposizione dei dati alle singole funzioni.
7. *Adozione di tool avanzati di analisi dei dati*: grado di utilizzo di strumenti a supporto di analisi previsionali che consentono di prevedere il valore futuro di variabili numeriche e

categoriche. È molto importante che le funzioni aziendali ricorrano a queste soluzioni e non solo a quelle tradizionali di Descriptive Analytics perché sono le prime che permettono di gestire con maggiore efficacia i processi decisionali.

8. *Numero di tipologie di soluzioni adottate per l'analisi dei dati*: quantità di classi diverse di strumenti per l'analisi dei dati adottate all'interno della funzione che ci consente di valutare il livello di pervasività delle soluzioni di BDA in questa.
9. *Grado di adeguatezza delle soluzioni adottate*: livello di supporto fornito dagli strumenti di Big DataAnalytics alla funzione aziendale.

Questi possono essere raggruppati in 3 categorie diverse: conoscenza e stato di diffusione attuale e futuro dei BDA (KPI 1, 2, 3), dati analizzati (KPI 4, 5, 6) e soluzioni adottate (KPI 7, 8, 9).

Ad ogni indicatore per ciascuna funzione aziendale attribuiamo un livello su una scala da 1 a 5, che rappresentiamo con uno specifico colore:

1. Livello 5 (verde scuro): la funzione aziendale si trova in uno stadio avanzato per quel KPI e può quindi trarne enorme vantaggio rispetto alle altre.
2. Livello 4 (verde chiaro): la posizione della funzione aziendale per quell'indicatore è buona, anche se c'è ancora spazio di miglioramento per arrivare alla fase più evoluta.
3. Livello 3 (giallo): la funzione aziendale si trova a metà del guado, quindi ha compiuto dei progressi nell'area definita dall'indicatore, ma non si trova ancora in una fase matura.
4. Livello 2 (arancione): la funzione aziendale sta iniziando ad agire per incrementare le prestazioni di quell'indicatore ma si trova in una fase ancora arretrata.
5. Livello 1 (rosso): la situazione della funzione aziendale per quell'indicatore è critica e quindi c'è ancora molto lavoro da fare per migliorarsi in quel KPI.

Dopo aver attribuito per ogni funzione aziendale i diversi livelli ai KPI, abbiamo quindi valutato lo stato di maturità di ciascuna, facendo una media dei livelli assegnati a tutti gli indicatori. Il risultato finale è riportato nella tabella nella pagina seguente.

Da quest'analisi emerge che è la funzione di *Marketing e Vendite* a presentare il più alto livello di maturità (4/5) nei Big DataAnalytics e ad essere quindi pronta a cogliere le grandi opportunità offerte da queste soluzioni, così come l'*Amministrazione, Finanza e Controllo* che mostra un livello medio/alto (3,3/5). Gli *Acquisti* e la *Logistica e Produzione* si trovano invece

KPI Funzioni aziendali	CONOSCENZA STATO DI DIFFUSIONE ATTUALE E FUTURO DEI BDA			DATI ANALIZZATI			SOLUZIONI ADOTTATE			LIVELLO MATURITÀ
	Livello di fruizione attuale	Potenzialità future	Consapevolezza dell'impatto strategico dei Big Data	Utilizzo di dati semi/destrutturati	Utilizzo di dati esterni	Qualità della gestione dei dati	Adozione di tool avanzati	Numero di tipologie di soluzioni adottate	Grado di adeguatezza delle soluzioni	
Marketing e Vendite										4/5
Amministrazione, Finanza e Controllo										3,3/5
Logistica e Produzione										2,8/5
Ricerca e Sviluppo										2,1/5
Risorse Umane										1,9/5
Acquisti										2,8/5

Tabella 8 : Cruscotto indicatori per la valutazione del livello di maturità delle funzioni aziendali

in uno stato intermedio (2,8/5), la *Ricerca e Sviluppo* in uno medio/basso (2,1), mentre le *Risorse Umane* risultano essere la funzione più arretrata in quest'ambito (1,9).

### **Marketing e Vendite (4/5)**

Cerchiamo ora di capire le ragioni alla base della posizione occupata dalla funzione di *Marketing e Vendite*. Innanzitutto quest'unità non solo è quella che fruisce maggiormente dei sistemi di BDA, come dichiarato dal 77% dei CIO intervistati, ma anche quella per cui si prevede una crescita vigorosa e continua: il 75% dei CIO reputa infatti che vi siano enormi potenzialità per il futuro. Al suo interno c'è inoltre grande coscienza del fatto che i Big Data rappresentino un elemento rivoluzionario che porterà allo stravolgimento di logiche competitive e a business model innovativi e questo è un aspetto che influenza molto il fatto di essere pronti ad adottare e ottenere vantaggi dal loro sfruttamento. Considerando invece il livello di utilizzo di dati semi/destrutturati possiamo osservare come la funzione si trovi ancora in una fase arretrata, dedicandosi maggiormente allo studio di dati strutturati.

Questa è un'area critica per tutte le unità, ad eccezione della *Ricerca e Sviluppo* che, nonostante il minor livello di maturità, sembra mixare maggiormente le due classi. Scendendo nel dettaglio delle tipologie di dati analizzati dagli Analytics, è possibile confermare tale risultato dalle informazioni della survey rivolta ai responsabili C-Level delle funzioni aziendali, in quanto per la grande maggioranza si tratta di dati transazionali, record e documentazioni office, quindi di natura strutturata. Altri dati analizzati appartenenti a questa categoria, ma in misura nettamente inferiore, sono i log Data, i dati di localizzazione e GPS, quelli M2M generati dai sensori (RFID, Bluetooth, NFC,...) e quelli originati da banche dati open. I dati di natura non strutturata presi in considerazione, seppure in quantità di gran lunga inferiore rispetto a quelli transazionali, sono email e file di testo, dati provenienti dai Social Media, seguiti da immagini, video e audio.

Tuttavia possiamo notare come la funzione di Marketing e Vendite si contraddistingua per un utilizzo maggiore dei dati esterni rispetto alla maggior parte delle altre unità, più tradizionali, le quali si dedicano per la grande maggioranza allo studio dei dati di natura interna, senza fare un'adeguata integrazione. Sono i dati social, sempre più importanti nel mondo attuale, ad essere analizzati in misura crescente (dal 43% degli indagati). Questo aspetto mette in evidenza l'importanza del fenomeno per la definizione delle strategie di Marketing, in quanto l'ascolto degli utenti e l'estrazione di insight relativi alle loro opinioni, alle loro preferenze e ai loro bisogni può portare ad enormi vantaggi.

Dati che invece non ricevono, ma dovrebbero ricevere, un'adeguata considerazione dalla funzione sono le immagini (solo il 22% li analizza): la loro quantità è in continua crescita visto i milioni di foto postate ogni giorno sui social media, la cui analisi sarebbe in realtà molto utile per individuare come i consumatori associano il brand e i prodotti dell'azienda ai vari contesti. Poco utilizzati sono anche i dati di localizzazione e GPS (10%), che rappresentano invece un enorme potenziale per il Marketing: conoscere la posizione del cliente per fare azioni di advertising in real time è un esempio.

Relativamente alla tipologia di dati utilizzati quindi, i risultati raccolti dalle survey e dalle interviste mettono in evidenza l'incapacità generale delle funzioni di sfruttare il potenziale del patrimonio a disposizione: tutte si limitano per lo più all'analisi dei dati transazionali, trascurando tutti gli altri. Si tratta di un errore enorme considerando il volume in continuo aumento di tutte le categorie di dati, specialmente di quelli dai social media. L'unica funzione che, come appena spiegato, sembra aver fatto dei passi in avanti è il Marketing.

Focalizzandoci ancora su quest'ultima, alla qualità dei dati abbiamo assegnato il livello 4. Non il 5 perchè parte delle unità di Marketing e Vendite indagate ha indicato di trovarsi ancora in una situazione tradizionale, in cui esistono policy e linee guida per gestire la qualità dell'informazione con un livello di affidabilità basso e standardizzato. Tuttavia la maggioranza raccoglie i dati a livello centralizzato e li mette a disposizione delle diverse applicazioni che ne richiedono l'uso con una qualità e affidabilità generalmente elevate e una parte, anche se minore, si trova in una situazione avanzata in cui i dati sono sempre disponibili aggiornati ed univocamente identificati.

L'ambito in cui la funzione occupa una posizione migliore rispetto alle altre è quella relativa alle soluzioni adottate, solo la Logistica e Produzione tiene il suo passo. Innanzitutto al suo interno vengono adottate più di una tipologia di strumenti, anche se per lo più si tratta di tool tradizionali e quindi di visualizzazione dei dati e delle informazioni e di produzione di reportistica, che effettuano analisi passive, mirando a rappresentare i dati mediante funzionalità di query e reporting, limitandosi solamente ad offrire una vista logica dell'esistente. Tuttavia l'adozione di tool a supporto di analisi previsionali è buona ed è destinata a crescere: essi sono molto importanti perché permettono per esempio di prevedere i bisogni dei clienti e i loro comportamenti al fine di identificare quelli che saranno più profittevoli o quelli che risponderanno con maggior probabilità ad una data campagna di Marketing. Il supporto da loro fornito è sufficiente, anche se non pienamente efficace, ma maggiormente adeguato alle esigenze dell'unità se confrontato a quello delle altre funzioni.

Queste sono le motivazioni che spiegano il maggior livello di maturità del Marketing e Vendite. La funzione si sta muovendo in una direzione che gli permetterà di cogliere sempre di più le opportunità offerte dai Big Data Analytics, come testimoniato per esempio dalla volontà di integrare in misure maggiore dati non convenzionali, quindi destrutturati ed esterni.

### **Amministrazione, Finanza e Controllo (3,3/5)**

L'*Amministrazione, Finanza e Controllo*, come già detto, si caratterizza per un livello di maturità medio/alto: nell'ambito dei Big Data Analytics dei passi in avanti sono già stati fatti, tuttavia c'è ancora molto lavoro da fare per migliorare. Innanzitutto insieme al Marketing e Vendite, questa funzione è il principale fruitore delle soluzioni di BDA all'interno delle organizzazioni, così come dichiarato dal 76% dei CIO indagati, tuttavia ha una prospettiva di crescita più bassa della prima, ma comunque superiore rispetto a quella delle altre funzioni. Essa ha infatti a disposizione soluzioni mature per navigare tra i KPI che consentono di capire come si sta muovendo l'azienda, le quali sembrano rappresentare però il passato e il presente, più che il futuro. Al suo interno c'è comunque coscienza dell'enorme impatto strategico e globale dei Big Data sull'intera organizzazione.

I valori positivi assunti da questi 3 KPI, insieme a quello relativo alla qualità dei dati, dove la funzione si trova in uno stadio avanzato dato che per lo più i dati raccolti sono aggiornati con un livello di affidabilità elevato, sono quelli che impattano positivamente sul suo stato di maturità. L'ambito relativo alla tipologia di dati analizzati è invece quello in cui si trova in uno stadio arretrato: l'integrazione tra dati di natura strutturata e non e di origine interna ed esterna manca, in quanto vengono considerati solo quelli tradizionali, in prevalenza dati transazionali seguiti con molto distacco da dati M2M generati dai sensori e dai misuratori digitali e dagli Open Data.

L'altra area in cui le prestazioni dell'Amministrazione, Finanza e Controllo non sono le migliori riguarda le soluzioni implementate: se da un lato vengono adottati diversi strumenti per l'analisi dei dati, solitamente si tratta di tool classici, di visualizzazione o reporting, e meno di predictive Analytics per prevedere per esempio indicatori di prestazione o il rischio finanziario, con un livello di supporto insufficiente.

### **Acquisti (2,8/5)**

Per la funzione *Acquisti* abbiamo individuato uno stato di maturità verso i Big Data Analytics intermedio. Si tratta di una buona adottatrice delle soluzioni di BDA secondo quanto affermato



dai CIO, il 55% dei quali l'ha indicata come funzione che fruisce di tali strumenti. Tuttavia la prospettiva di crescita futura in quest'ambito è abbastanza bassa, al contrario di quello che ci saremmo aspettati, visto che un eventuale sfruttamento dei Big Data potrebbe portare a raccogliere benefici in diverse attività, tra cui quelle di spend analysis, di sviluppo della strategia di acquisto e di valutazione dei fornitori. Il fatto che all'interno della funzione ci sia consapevolezza del ruolo strategico che i Big Data avranno è in contrasto con l'assenza prevista di opportunità future.

Dati non strutturati e dati esterni non ricevono la giusta attenzione, anche se per esempio quelli provenienti da fonti esterne, come il web, i social media e i sensori potrebbero essere utilizzati per la selezione dei fornitori, per la negoziazione e per un'anticipazione del vendor rating e quindi per la valutazione delle performance dei fornitori. Non ci sono invece problemi per quanto riguarda la qualità dei dati.

Un elemento critico, che pesa molto nella valutazione, è invece rappresentato dal supporto insufficiente e non adeguato alle esigenze fornito dalle soluzioni adottate, che sono di diversi tipi, per lo più di visualizzazione e reportistica, anche se i tool previsionali hanno iniziato ad essere implementati e questo è un elemento che gioca a favore della funzione.

### **Logistica e Produzione (2,8/5)**

La *Logistica e Produzione* presenta un livello medio di maturità, così come gli Acquisti, anche se le loro aree di forza e di debolezza sono diverse. All'interno delle organizzazioni questa è una funzione che fruisce mediamente dei BDA e per la quale i CIO indagati non prevedono enormi opportunità future. Quest'ultimo è un risultato che disattende un po' le nostre attese in quanto tali soluzioni potrebbero essere utilizzate per la previsione della domanda con l'obiettivo di ottimizzare la capacità di trasporto e l'allocazione del personale, per l'ottimizzazione in real time dei percorsi e per la previsione dei rischi di supply chain, con conseguenti vantaggi in termini di costo. Anche all'interno dell'unità non sono ancora state comprese al massimo le opportunità offerte dai BDA.

In linea con le altre funzioni, non viene analizzato un giusto mix di dati strutturati e non e di dati interni ed esterni, ma c'è una predilezione allo studio dei primi e dei terzi. Per la Logistica e Produzione potrebbe essere utile ricorrere ai dati provenienti dai Social Media per rilevare tempestivamente dati sugli sviluppi politici locali, sugli eventi naturali e meteorologici e quindi su tutte le fonti di rischio in modo da individuare un'eventuale situazione di criticità per la supply chain del cliente. Tuttavia questo non viene ancora fatto. La qualità dei dati è buona: la

loro raccolta a livello centralizzato e la loro messa a disposizione a chi ne ha bisogno è affidabile. In realtà questa è una dimensione in cui tutte le funzioni si trovano in una buona posizione, vicine alla piena maturità.

È relativamente alle soluzioni adottate che la funzione ha fatto dei passi in avanti, più dell'Amministrazione, Finanza e Controllo e degli Acquisti: nella maggior parte viene utilizzato più di uno strumento, sebbene si tratti per lo più di Basic Analytics, anche se possiamo osservare una prima implementazione di predictive Analytics e il supporto fornito è sufficiente.

### **Ricerca e Sviluppo (2,1/5)**

Per la *Ricerca e Sviluppo* c'è molto lavoro da fare nell'ambito dei BDA considerato il suo livello di maturità medio/basso. Osservando in primis lo stato di diffusione attuale e futuro di queste soluzioni, la funzione sembra essere nella posizione peggiore rispetto a tutte le altre, in quanto non solo è quella che usufruisce in misura minore degli strumenti di BDA, ma è anche quella, insieme alle Risorse Umane, per cui si prevedono le minori potenzialità, come dichiarato dai CIO. In realtà queste soluzioni potrebbero essere utilizzate per l'identificazione dei bisogni dei clienti per nuovi prodotti/servizi, per il loro design e per la raccolta continua di feedback su di questi. L'area relativa alla tipologia di dati utilizzati sembra essere quella in cui la funzione è messa meglio, utilizzando più di tutti, anche del più maturo Marketing e Vendite dati non strutturati ed esterni. Tuttavia questo risultato ci sembra un po' strano, considerando soprattutto il basso livello di utilizzo dei BDA, e molto probabilmente è dovuto al campione indagato di C-level di Ricerca e Sviluppo poco numeroso che ha portato ad un esito a nostro avviso distorto. Infine per quanto riguarda l'area delle soluzioni, la funzione appare ancora arretrata rispetto alle altre. La funzione quindi si trova ancora in una fase iniziale.

### **Risorse Umane (1,9/5)**

Come già accennato, le *Risorse Umane* non risultano pronte ad adottare a pieno i BDA, essendo caratterizzate da un basso livello di maturità. Attualmente la funzione non sfrutta abbastanza le soluzioni: solo il 31% dei CIO intervistati l'ha indicata come fruitrice di tali strumenti e solo il 22% prevede potenzialità future. Questo risultato è in contrasto con le reali opportunità offerte in quanto i BDA potrebbero essere sfruttati nella valutazione delle performance, nella definizione delle politiche retributive, nella gestione della forza lavoro e della formazione e ad un livello più avanzato nella previsione del tasso di abbandono, delle prestazioni, dell'assenteismo e del successo dei candidati. Quest'ultimo ambito mette quindi in evidenza il potenziale messo a disposizione dai tool previsionali, i quali tuttavia vengono scarsamente

adottati dalla funzione, la quale si serve per lo più di una sola tipologia di soluzione di Basic Analytics, con un supporto fornito non adeguato alle esigenze. Le diverse classi di dati non vengono mixate, mentre, in linea con altre unità, la qualità della raccolta e della gestione dei dati è buona. Questo quadro poco roseo può essere in parte spiegato dal fatto che le Risorse Umane non abbiano ancora compreso le potenzialità che i sistemi di BDA offrono.

Lo stato di maturità identificato per le varie funzioni aziendali le quali, chi più chi meno, non riescono ancora a cogliere al 100% le opportunità offerte dai sistemi di BDA, può essere giustificato dalla serie di ostacoli che si trovano ad affrontare, condivisi da tutte. Tali barriere possono essere raggruppate in diverse categorie: *tecniche*, legate all'integrazione e alla qualità dei dati, *tecnologiche*, relative al livello di efficacia delle soluzioni adottate, *organizzative/gestionali*, riguardanti la gestione dei progetti Big Data all'interno delle funzioni, *legate alle competenze*, relative alla presenza di figure adeguate per queste iniziative, *culturali*, correlate all'apertura verso i Big Data, *economiche* e *legate alle privacy*. Dall'analisi delle risposte delle diverse funzioni emerge che la maggior parte delle sfide da sostenere sono legate alle competenze, seguite da quelle di natura culturale ed economica, tecnica e organizzativa/gestionale; non risulta invece un problema la questione tecnologica.

Scendendo maggiormente nel dettaglio, il principale freno è rappresentato dalla mancanza di figure specializzate quali il Data Scientist e lo Chief Data Officer (*legate alle competenze*), altre barriere importanti sono la difficoltà di integrazione con altri applicativi (tecnica), la complessità nello stimare il ritorno dell'investimento (*legate alle competenze*), gli elevati tempi di implementazione e la mancanza di coinvolgimento del top management (*organizzative/gestionali*) e l'ingente investimento (economica).

### 5.3 IL RUOLO DELLE STARTUP

Un ruolo essenziale nell'ambito Big Data Analytics & Business Intelligence è ricoperto dalle startup che offrono alle aziende l'opportunità di incrementare l'innovazione sia dei processi interni che di quelli di front-end. Le soluzioni proposte possono quindi portare a vantaggi quali, da un lato, l'ottimizzazione dei processi produttivi e logistici o il miglioramento di quelli di selezione e formazione HR e dall'altro lato l'incremento dell'efficacia delle campagne di Marketing e una maggiore customer satisfaction.

In collaborazione con l'Osservatorio Startup Hi-tech del Politecnico di Milano abbiamo effettuato una ricerca sulle startup che offrono soluzioni di BDA&BI. In particolare abbiamo analizzato a livello internazionale 498 startup, di cui 33 italiane, che dal 2012 hanno ricevuto finanziamenti pari a 14.476 milioni di dollari da investitori istituzionali.

### 5.3.1 DATI ANAGRAFICI

Relativamente alla distribuzione geografica, il 77% delle startup considerate proviene dal Nord America, con un netta dominanza degli Stati Uniti (98%). Il restante 23% si suddivide tra Europa (17%), Asia (4%) di cui circa la metà sono israeliane, Sudamerica (2%) e Africa (meno dell'1%); si tratta comunque di percentuali irrisorie rispetto a quella nordamericana, che mettono quindi in evidenza come in questa zona il livello di innovazione in quest'ambito abbia raggiunto livelli molti avanzati.

La maggior parte delle startup è stata fondata dal 2009 (70%) con gli anni dal 2010 al 2013 inclusi caratterizzati dalla maggior prolificità (56%), in quanto è in questo periodo che le aziende hanno compreso realmente l'importanza di sfruttare i dati a loro disposizione. Possiamo tuttavia osservare un calo delle nascite dal 2010 al 2011 legato alla crisi mondiale, seguito però da un recupero nell'anno successivo e da un nuovo decremento dovuto alla scarsa pubblicizzazione delle startup stesse. Quelle nate nel 2005 sono invece le imprese che hanno intrapreso per prime le sfide della BI e sono state in grado di evolversi nel tempo.

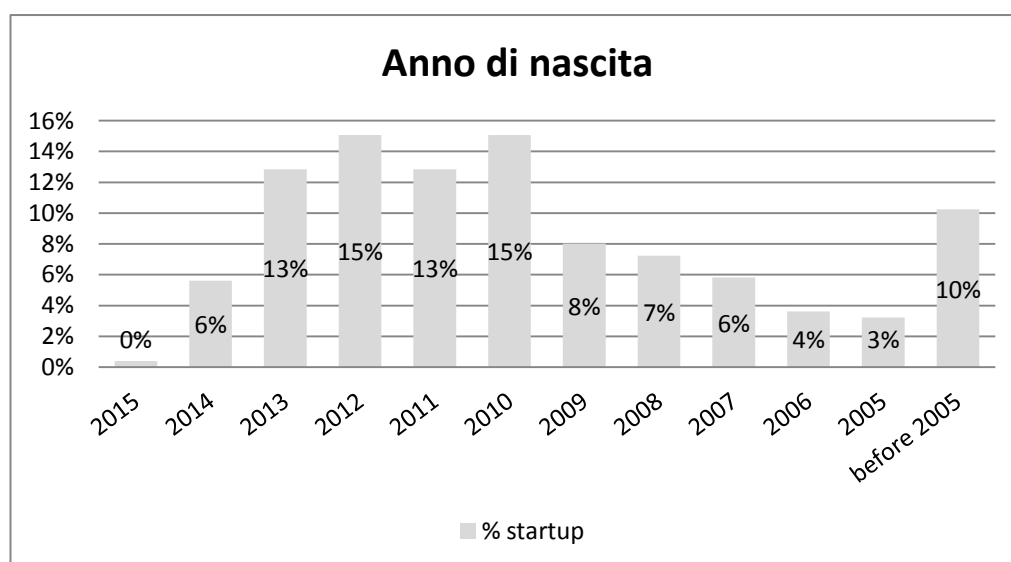


Grafico 11: Scomposizione delle startup internazionali per anno di nascita

Le startup censite offrono per la grande maggioranza applicazioni “general purpose”, orientate quindi a diverse industry (75%), solo il 25% è focalizzato su uno specifico settore, per lo più il Retail (7%) e l’Health (5%).

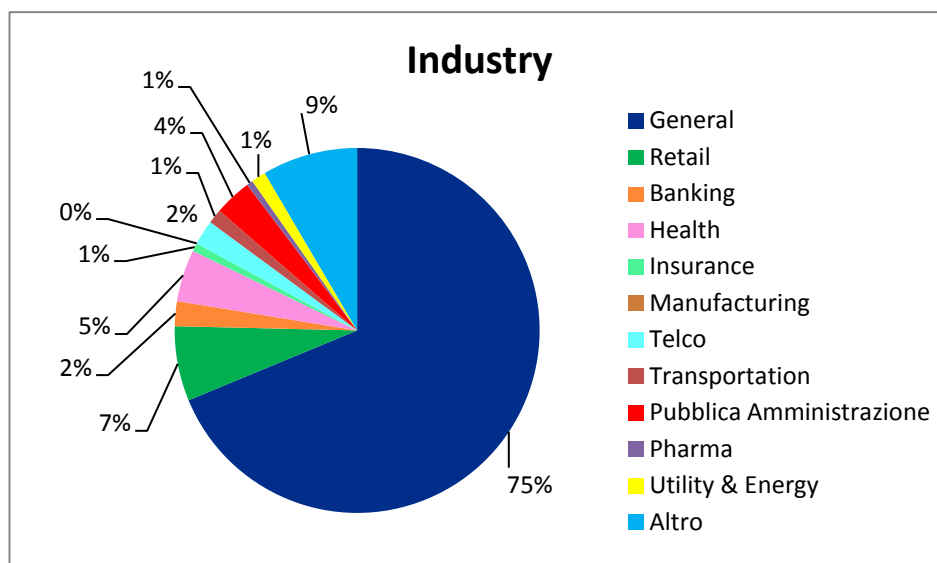


Grafico 12: Scomposizione delle startup internazionali per industry

Business Angel, Venture Capitalist e società di investimento hanno finanziato le startup per un ammontare pari a 18.300 milioni di dollari complessivamente, di cui 14.476 dal 2012. In particolare nell’ultimo anno si è verificato un incremento notevole, pari a 6.856 milioni di dollari. Suddividendo i finanziamenti in fasce possiamo notare dal grafico come la maggior parte delle startup ne abbia ricevuti meno di 2 milioni (25%) e tra i 20 e i 50 esclusi (21%), il 15% ne ha ottenuti tra i 10 e i 20 esclusi, il 14% tra i 5 e i 10 esclusi. Da notare come 3 startup

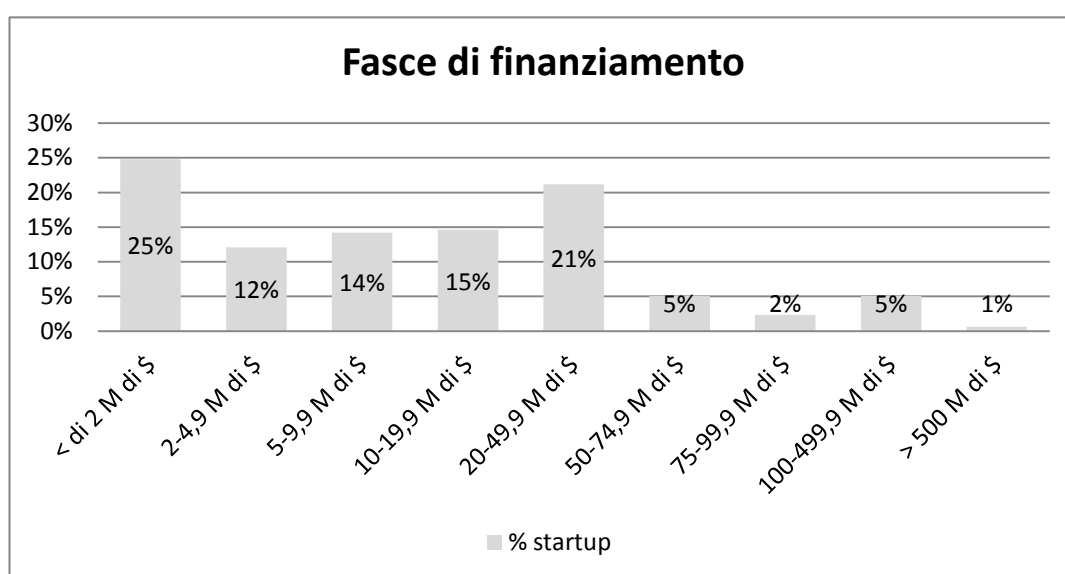


Grafico 13: Suddivisione delle startup internazionali in base finanziamenti ricevuti dalle dal 2012

siano state finanziate per più di 500 milioni di dollari: si tratta di Cloudera (1.125 milioni), Palantir Technologies (1.102 milioni) e BrightBytes (514 milioni).

### 5.3.2 DIMENSIONI DI ANALISI

Noi possiamo classificare le startup secondo tre diverse dimensioni di analisi:

- *tipologia dei dati*: strutturati o non strutturati. I primi sono quelli caratterizzati da uno schema, quindi di fatto quelli gestiti dai DBMS classici con alte prestazioni, mentre i secondi sono dati completamente privi di schema, i quali possono essere distinti in destrutturati come immagini, testi, video e semistrutturati, per i quali esiste una struttura parziale, nonsufficiente tuttavia per permetterne la memorizzazione e la gestione da parte dei Database relazionali.
- *Origine dei dati*: interni o esterni. I primi vengono generati dai processi interni all'azienda, i secondi da processi esterni, come i dati dai social media o quelli relativi al contesto di riferimento.
- *Ambiti applicativi*

Per quanto riguarda la prima dimensione di analisi, la Ricerca mette in evidenza come la maggior parte delle soluzioni offerte dalle startup analizzi sia dati strutturati che dati non strutturati (62%) e come una percentuale molto più bassa studi esclusivamente dati o solo strutturati (22%) o solo non strutturati (16%).

Relativamente alla loro origine risulta che il 27% delle startup propone applicazioni che

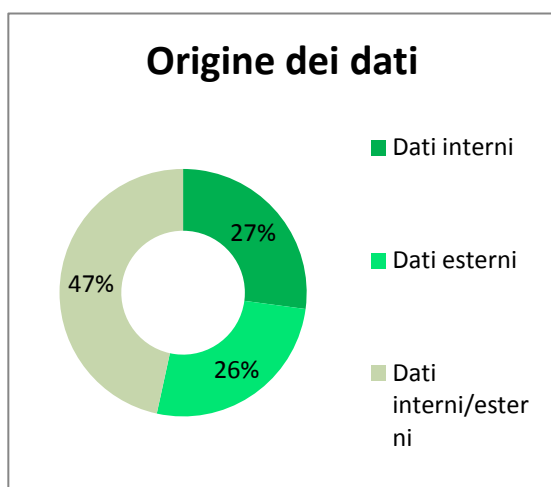


Grafico14: Scomposizione delle startup internazionali per origine di dati analizzati

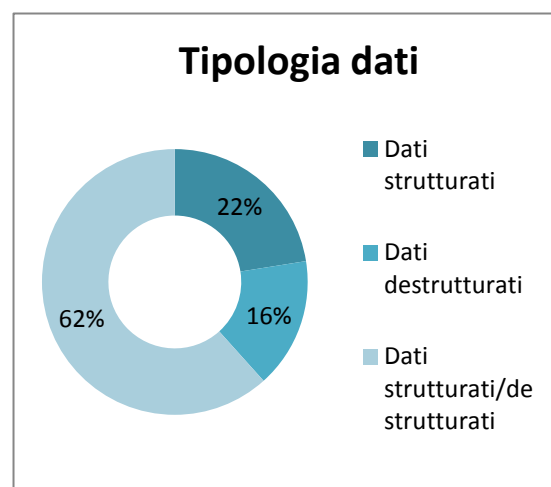


Grafico 15: Scomposizione delle startup internazionali per tipologia di dati considerati

analizzano solo dati interni, il 26% solo dati esterni, mentre la maggioranza, il 47%, entrambi. Andando a confrontare i risultati dell'indagine dell'anno scorso e quelli dell'indagine dell'anno corrente, che ha aggiunto alla lista altre 122 startup, non si osservano particolari scostamenti.

Nello specifico, andando ad incrociare le due dimensioni, possiamo osservare come la maggior parte offra soluzioni che esaminano dati strutturati e destrutturati provenienti da fonti sia interne che esterne (35%). Questo è un risultato che deve attirare la nostra attenzione, in quanto mette in risalto la possibilità Data alla aziende di mixare un numero maggiore di dati diversi e di estrarre quindi insight più significativi e di conseguenza maggiore valore. Queste startup sono poi seguite da quelle che dedicano la loro attenzione a dati interni di entrambi i tipi (14%) e da quelle che svolgono analisi tradizionali, considerando quindi solo dati di natura strutturata e di origine interna (13%).

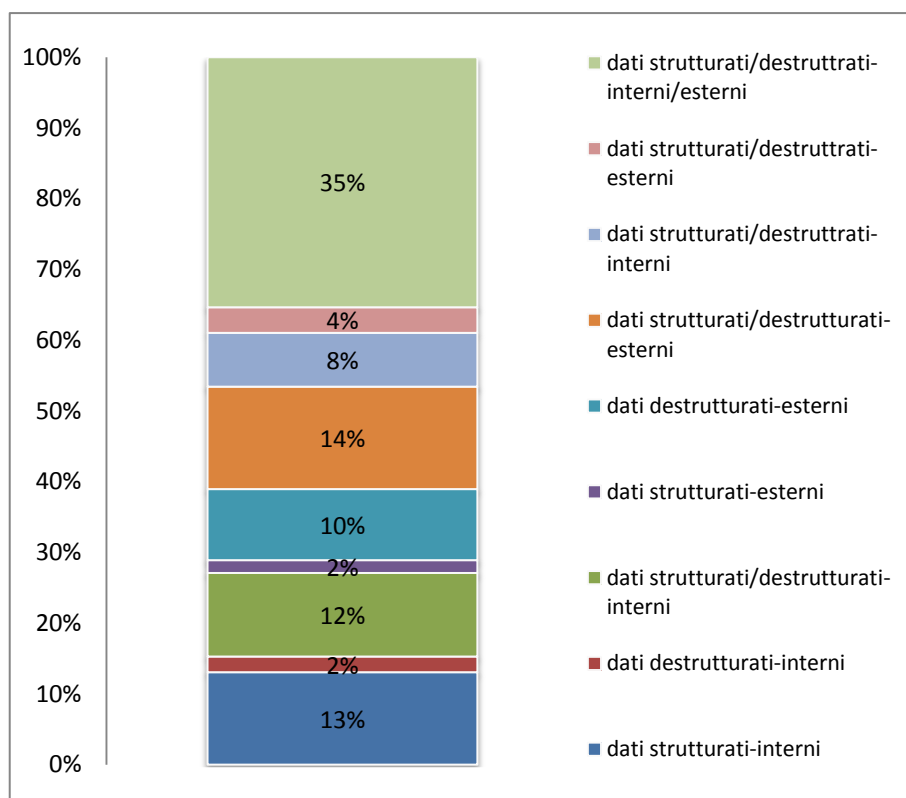


Grafico 16: Scomposizione delle startup internazionali per tipologia e origine di dati analizzati

Considerano infine gli ambiti applicativi, le startup possono classificate nelle tre seguenti categorie:

1. *Infrastructure*: startup che offrono infrastrutture che processano, memorizzano e analizzano i dati, rappresentando il cuore dell'ecosistema Big Data. Al loro interno è possibile distinguere quattro sottocategorie:
  - *Database* (24%): mettono a disposizione sistemi software progettati per consentire la creazione, la manipolazione e l'interrogazione efficiente dei dati. Si tratta di Database NoSQL in grado di gestire una grande varietà di dati strutturati, semi-strutturati e destrutturati, più scalabili e dalle performance migliori (MongoDB, Couchbase), di Database NewSQL che cercano di fornire le stesse prestazioni scalabili dei sistemi NoSQL per l'elaborazione delle transazioni online in lettura e scrittura (Clustrix, NuoDB) e di MPP Database che lavorano segmentando su più nodi i dati i quali vengono elaborati in parallelo (ParAccel).
  - *Technologies* (50%): offrono tecnologie e linguaggi grazie alle quali è possibile effettuare analisi Big Data (Cloudera, Hortonworks, MapR).
  - *Storage* (6%): propongono soluzioni di storage altamente performanti, in grado di gestire migliaia di terabyte di dati, utilizzando la potenza in modo più efficiente con conseguenti costi di storage minori rispetto alle tecnologie tradizionali (AmpliData, Argil Data Corp.).
  - *Cross-infrastructure* (20%): offrono infrastrutture ma di carattere generale (Moogsoft, Delphix).
  
2. *Analytics*: startup che mettono a disposizione specifiche tecnologie aventi capacità analitiche. Anche per questa categoria si possono individuare cinque classi di soluzioni offerte:
  - *Data visualization* (24%): piattaforme che prendono i dati grezzi e li presentano in formati visivi complessi e multidimensionali per rappresentare graficamente i principali KPI e mettere in risalto l'informazione (SiSense, Perceivant, Quid, Visual.ly).
  - *Search* (3%): strumenti che consentono di esplorare i dati tramite query, interpretate dal motore di ricerca che analizza le informazioni e restituisce velocemente le risposte. Si tratta di sistemi di analisi semantica che apprendono dalla sintesi e dal processamento dei dati (Widzee, Recommind).
  - *Location/people/events* (4%): Analytics che forniscono informazioni geografiche per identificare la localizzazione di un utente mobile o web sfruttando tecniche quali il GPS o le reti WiFi e WLAN, per esempio per visualizzare i dati su mappe



personalizzate al fine di supportare algoritmi di clustering o per ricavare insight relativi ai comportamenti dei clienti basandosi sulla loro posizione (Digby, Fliptop, Radius).

- *Analytcs platform* (47%): piattaforme che integrano e analizzano i dati per estrarre insight in tempi rapidi, aiutando le aziende a prendere decisioni migliori (Data Systems, Guavus, HydroPoint).
- *Social Analytcs* (27%): soluzioni che mettono a disposizione tool di Behavioral & Cohort Analysis, Sentiment Analysis, Brand Reputation e identificazione degli influencer sui social networks (Bitly, Dataminr, DataSift, HootSuite,).

3. *Application*: startup che offrono soluzioni verticali di analisi di dati e molto spesso di piattaforme tecnologiche abilitanti rivolte a uno o più dei seguenti ambiti applicativi:

- *Marketing& Sales* (50%): per la gestione delle vendite, per l'incremento dell'efficacia delle campagne pubblicitarie sui diversi canali e in particolare del Digital Marketing, per la personalizzazione dei messaggi di Marketing e per l'ottimizzazione dei prezzi (Lattice Engines, Metamarkets, Rocket Fuel).
- *CRM & Customer Experience* (47%): per la gestione della relazione con i clienti e dei reclami (Pursway, Reflektion) per lo studio e il miglioramento dell'engagement del cliente, della Customer Experience e quindi del tasso di conversione (ClickFox, ClickTale), e anche dell'E-commerce (Jirafe, Minubo).
- *Finance* (15%): per l'analisi e il miglioramento delle performance finanziarie, della management strategy, della Governance e della gestione del rischio (Kreditech, Lendup, Ondeck).
- *Supply Chain, Logistica & Operations* (8%): per l'analisi e l'ottimizzazione dei costi logistici, del processo produttivo, del magazzino, del trasporto e del rapporto con i fornitori (Antuit, Decysion).
- *IT* (3%): per l'aumento dell'efficienza delle operazioni IT e per l'identificazione e la risoluzione rapida dei problemi (Splunk).
- *Human Resource* (6%): per l'analisi della forza lavoro e per l'ottimizzazione e il miglioramento dei processi di selezione, assunzione, formazione, valutazione delle performance e gestione delle paghe e delle ricompense (Entelo, Evolv, Gild).
- *Security* (15%): per l'analisi e l'incremento della sicurezza dei dati (Cyphort, Endgame, SkyHigh Networks).

Dalla ricerca emerge che le soluzioni maggiormente proposte dalle startup nella prima categoria sono le Technologies, mentre nella seconda gli Analytics platform e i Social Analytics, a testimonianza della volontà delle aziende di sfruttare i dati dei vari Facebook, Twitter, Instagram per trarre vantaggi competitivi considerando un mondo sempre più social. Rapportando i risultati della Ricerca di quest'anno con quelli dell'anno precedente, possiamo osservare un incremento nell'adozione di soluzioni Social Analytics, passata dal 17% al 27%, dato che conferma la rilevanza sempre maggiore del fenomeno social.

Infine, relativamente all'ultima categoria, le applicazioni proposte in misura maggiore sono quelle di Sales & Marketing e di CRM & Customer Experience, quindi una prevalenza di quelle rivolte alla relazione con il cliente piuttosto che ai processi interni.

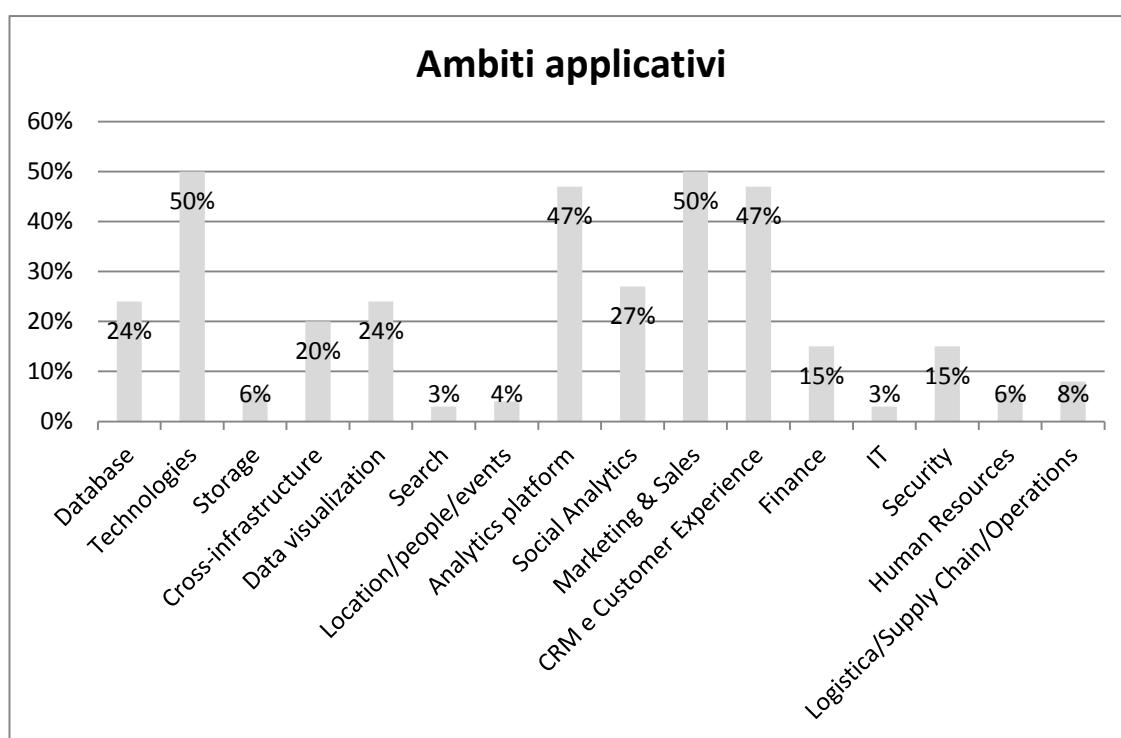
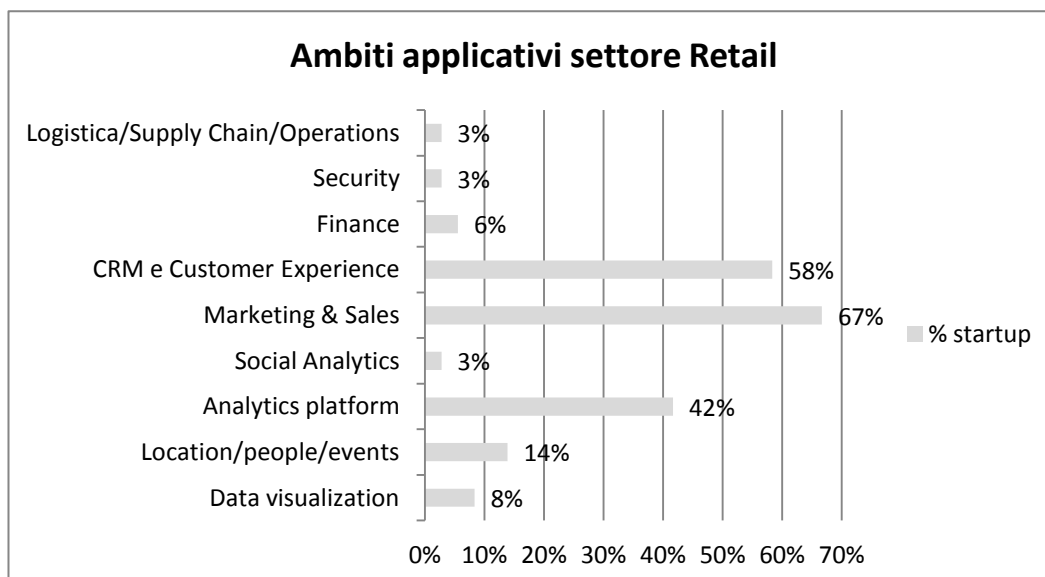


Grafico 17: Scomposizione delle startup internazionali per ambiti applicativi

Concentriamoci ora sul Retail, che abbiamo visto essere il settore su cui si focalizzano principalmente le soluzioni delle startup. Abbiamo già visto nel primo capitolo come i Big Data possano essere sfruttati per trarre dei vantaggi sia nei processi di back-end come quelli logistici e di gestione del magazzino, sia in quelli di front-end, relativi all'interazione con il consumatore. Dall'analisi dei dati emerge come le startup mettano a disposizione dei Retailer applicazioni per la seconda area: il 67% offre infatti soluzioni di Marketing & Sales e il 58% soluzioni di CRM & Customer Experience, mentre nel campione analizzato nessuno propone

quelle di Logistica e di Supply Chain. Questo risultato inizia a mettere in risalto come vi sia una maggiore consapevolezza delle opportunità messe a disposizione dalle soluzioni BDA&BI in ambito Marketing e come l'offerta sia quindi adeguata a questo trend per far fronte alle maggiori esigenze delle aziende.

Infine incrociando ambiti applicativi e finanziamenti ricevuti possiamo osservare come le startup per cui società di investimento, Business Angel, Venture Capitalist hanno maggiormente investito sono quelle che offrono Technologies e Database (finanziamento medio pari rispettivamente a 47.598 e 43.946 milioni di dollari). Possiamo però notare anche i rilevanti investimenti nelle applicazioni di Security e Finance, per un ammontare medio di 77.991 e 51.602 milioni di dollari.



*Grafico 18: Scomposizione delle startup internazionali focalizzate sul settore retail per ambiti applicativi*

### 5.3.3 LE STARTUP ITALIANE

Approfondiamo ora la situazione in Italia, in cui abbiamo individuato 33 startup molto giovani, il 58% di queste sono state infatti fondate dal 2013 e solo il 6% prima del 2010. Dall'analisi della loro distribuzione geografica emerge che è il Nord la zona dove si ha la maggior concentrazione (55%), seguita dal Centro (33%) e dal Sud e isole (12%). Confrontando questi risultati con quelli dell'anno scorso, in cui sono state analizzate 14 startup, è possibile confermare il dominio del Nord in termini di numero di startup situate e un ruolo più importante del Centro, in cui la percentuale è passata dal 22 al 33. Nello specifico è la Lombardia la regione con la maggior

concentrazione (36%), seguono l'Emilia Romagna (15%), il Veneto e il Lazio (12%); da sottolineare anche la posizione della Sicilia, in cui si trova il 9% delle startup.

A partire dal 2012 Business Angel, Venture Capitalist e società di investimento hanno finanziato le startup per un ammontare pari a 32, 5 milioni di dollari (abbiamo i valori solo di 18 aziende): in particolare sono quelle della Lombardia ad aver ricevuto i maggiori finanziamenti, pari a 21.1 milioni di dollari (65% del totale), tuttavia anche quelle del Sud ne hanno ottenuti di buoni, il 21% del totale (11% in Sicilia e 10% in Calabria). Si tratta comunque di cifre inferiori rispetto a quelle a livello internazionale: il 72% delle aziende ha ottenuto finanziamenti minori di 2 milioni di dollari, il 17% si trova nella fascia dai 2 ai 5 esclusi e l'11% in quella dai 5 ai 10 esclusi, nessuna ne ha ottenuti più di 10.

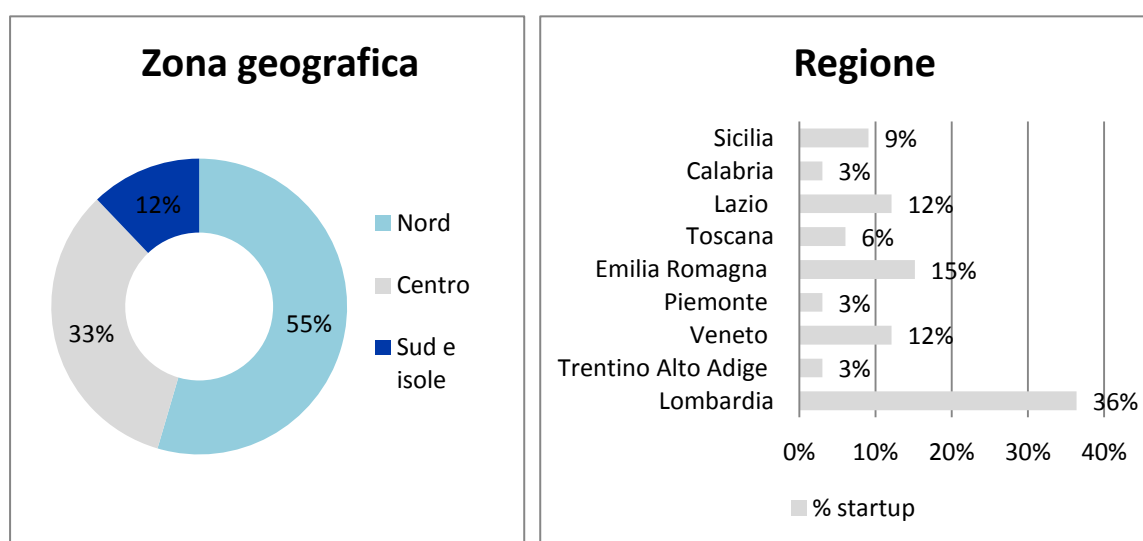
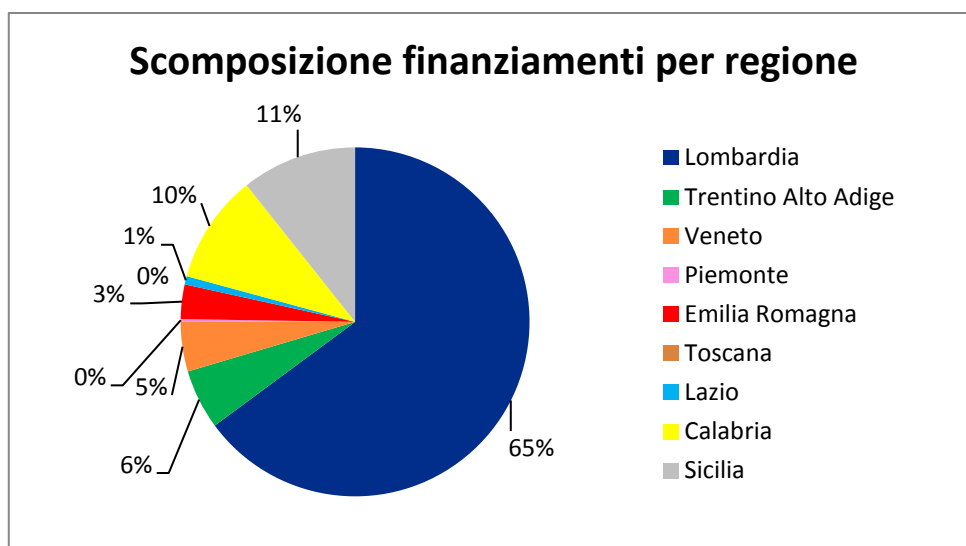


Grafico 19: Scomposizione delle startup italiane per zona geografica

Grafico 20: Scomposizione delle startup italiane per regione

In linea con i risultati a livello internazionale, le startup italiane offrono per lo più soluzioni generali, rivolte indistintamente ai diversi settori (61% del totale), la restante parte presenta una verticalizzazione forte, in particolare per il settore Retail (12%). Nel maggior numero di casi vengono analizzati contemporaneamente sia dati strutturati che non (19 su 33), tuttavia sono i dati esterni ad essere presi maggiormente in considerazione (15 su 33). Le soluzioni proposte sono per lo più di Data Visualization (27%), Social Analytics (24%) e Analytics Platform (21%) e le applicazioni per il Marketing & Sales (39%) e di CRM & Customer Experience (45%). Queste ultime due sono anche le soluzioni che ricevono più finanziamenti, insieme a quelle di Location/people/event e ai Social Analytics.



*Grafico 21: Suddivisione dei finanziamenti ricevuti dalle startup italiane a partire dal 2012 per regione*

Le tre startup maggiormente finanziate sono le seguenti:

- Beintoo: fondata nel 2011, crea un ecosistema per advertiser e publisher che rende possibile un nuovo approccio al digital advertising basato sui reali interessi degli utenti, in base ai quali questi vengono profilati. Beintoo ha ricevuto nel 2012 finanziamenti pari a 7 milioni di dollari in due turni.
- Cubeyou: nata nel 2010 a Milano, ha ottenuto dal 2012 finanziamenti pari a 4,06 milioni di dollari. Essa gestisce i dati social e analizza le relazioni tra gli interessi e i comportamenti dei consumatori per estrarre insight utili per la definizione di un'efficace strategia di Marketing e per lo sviluppo di campagne pubblicitarie che attirino i consumatori e aumentino il loro tasso di fidelizzazione.
- Admantax: fornisce dati e analisi di contesto, offrendo a publisher e advertiser soluzioni basate sui cookie che consentono di sviluppare un migliore advertising online. Admantax è stata fondata nel 2010, ricevendo negli ultimi 4 anni 5, 4 milioni di dollari di finanziamenti.

Abbiamo infine analizzato in dettaglio la figura del loro fondatore, considerando sia il profilo delle competenze che l'età. Per quanto riguarda il primo aspetto, a livello di startup, è emerso che nel 25% dei casi le competenze di chi ha fondato la startup sono tecniche, nel 16% manageriali e nel 59% sia tecniche che manageriali. Questo risultato è anche legato al fatto che circa la metà delle startup è stata fondata da più imprenditori (alcuni più tecnici altri più

manageriali), favorendo quindi la presenza di un bagaglio completo di competenze. Considerando invece il profilo del singolo, è risultato che la maggioranza ha competenze sia tecniche che manageriali, ma una percentuale comunque alta, il 40%, ne ha più tecniche e il restante 15% più manageriali. Relativamente all'età, nel 38% dei casi è compresa tra i 30 e i 40 anni, nel 33% è maggiore di 40 e nel restante 29% è compresa tra i 20 e i 30; quest'ultimo dato indica un crescente spirito di iniziativa e di innovazione da parte dei giovani.

Incrociando infine queste due dimensioni, possiamo osservare come le startup siano state fondate in prevalenza da persone dai 30 ai 40 anni con competenze tecniche (19%), seguite da quelle create da persone con abilità sia tecniche che manageriali di età compresa tra i 30 e i 40 o maggiore di 40 (15%).

		Competenze		
Età		Tecnico	Manageriale	Tecnico-Manageriale
	20-30	10%	6%	13%
	30-40	19%	4%	15%
	>40	12%	6%	15%

Tabella 9: Età e profilo di competenze dei fondatori delle startup italiane

Per concludere, facendo un confronto tra la situazione italiana e quella globale possiamo vedere come il nostro Paese si trovi in uno stadio più arretrato. Basta vedere il basso numero di startup censite pari a 33 e l'enorme differenza per quanto riguarda i finanziamenti: nessuna azienda italiana ha infatti ottenuto più di 10 milioni di dollari mentre nel resto del mondo ben 231 ne hanno ricevuti di più. A livello di industry, di tipologie di dati analizzati e di ambiti applicativi invece, come già detto, non ci sono differenze di risultato.

#### 5.4 SUPPORTO DELL'OFFERTA DI MERCATO ALLE FUNZIONI AZIENDALI NELL'AMBITO DEI BIG DATA ANALYTICS

Per verificare il supporto dell'offerta alle varie funzioni aziendali abbiamo identificato a quali di queste le soluzioni offerte dalle startup appena descritte si rivolgono. La nostra attenzione si è focalizzata su tali aziende per l'alto tasso di innovazione che le contraddistingue e questo è un elemento fondamentale in un mondo come quello dei Big Data, all'avanguardia e che richiede l'introduzione di continue novità. Riportiamo i risultati nella tabella nella pagina seguente, in

Soluzioni Offerte Funzioni aziendali		APPLICATIONS						
		% di diverse soluzioni offerte	69%	63%	56%	44%	63%	50%
	ANALYTICS	IT (3%)						
		Human Resource (6%)					X	
		Supply Chain, Logistica & Operations (8%)			X			
		Finance (15%)		X				X
		Security (15%)		X				
		CRM & Customer Experience (47%)	X				X	
		Marketing & Sales (50%)	X					
		Search (3%)	X	X	X	X	X	X
		Location/people/event (4%)	X					
		Data Visualization (24%)	X	X	X	X	X	X
	APPLICATIONS	Social Analytics (27%)	X	X	X		X	
		Analytics Platform (47%)	X	X	X	X	X	X
		Marketing e Vendite						
		Amministrazione, Finanza e Controllo						
		Logistica e Produzione						
	ANALYTICS	Ricerca e Sviluppo						
		Risorse Umane						
		Acquisti						

Tabella 10: Livello dell’offerta di mercato alle funzioni aziendali nell’ambito dei BDA

cui sulle righe abbiamo considerato le funzioni aziendali e sulle colonne le soluzioni messe a disposizione dalle startup. Incrociando le due dimensioni abbiamo individuato il sostegno dell'offerta a ciascuna di queste. L'entità dell'offerta di ciascuna soluzione, che abbiamo riportato nella tabella in termini percentuali, è stata ricavata dalla classificazione per ambiti applicativi delle startup, spiegata nel paragrafo precedente.

Dalla tabella emerge come le soluzioni offerte coprano bene le esigenze delle funzioni di Marketing e Vendite, di Amministrazione Finanza e Controllo e di Risorse Umane le quali possono usufruire rispettivamente del 69%, 63% e 63% delle soluzioni totali. Comunque, seppure in misura minore, anche alle altre viene messo a disposizione un buon set di strumenti. Se però consideriamo le percentuali di startup che offrono le diverse applicazioni, possiamo vedere come una netta prevalenza fornisca quelle di Marketing & Sales (50%) e di CRM & Customer Experience (47%), delle quali usufruisce per lo più la funzione di Marketing e Vendite. Dunque l'offerta a disposizione di quest'ultima è molto ampia. Da questo punto di vista anche l'Amministrazione Finanza e Controllo e le Risorse Umane possono contare su una buona ed è proprio a queste tre, oltre alla Logistica e Produzione, che le startup propongono anche soluzioni di Social Analytics, che rappresentano il 27% dell'offerta. Non abbiamo inserito le soluzioni infrastrutturali perché supportano tutte le funzioni aziendali, ma le abbiamo considerate ugualmente nel calcolo delle percentuali di diverse soluzioni offerte.

Abbiamo seguito un procedimento analogo per la domanda di soluzioni di Big Data Analytics da parte delle funzioni aziendali. Nello specifico, avendo a disposizione dalla survey CIO i dati sulla domanda delle soluzioni di BDA, abbiamo associato a ciascuna di queste le funzioni aziendali che le richiedono, al fine di identificare le unità caratterizzate da una richiesta maggiore di soluzioni diverse. La tabella nella pagina successiva, in cui abbiamo considerato sulle righe le funzioni aziendali e sulle colonne gli ambiti di adozione dei BDA in modo da individuare la domanda di ciascuna funzione, mostra i risultati. Riportiamo prima una breve descrizione degli ambiti in cui vengono richieste le soluzioni.



IT Operations						
Soluzioni verticali per il business	X	X	X	X	X	X
Transportation Analytics			X			
Top Manager Dashboard Solutions	X	X	X	X	X	X
Telecommunication Analytics	X					
Supply Chain Analytics			X			
Social & Web Analytics	X	X			X	
Security Analytics		X				
Production, Planning & Sales	X		X			X
Human Resources					X	
Finance & Accounting		X				X
e-commerce	X					X
Customer Experience Analytics	X					
CRM Analytics	X					
<b>Ambiti richiesti</b> <b>Funzioni Aziendali</b>	<b>Marketing e Vendite</b>	<b>Amministrazione, Finanza e Controllo</b>	<b>Logistica e Produzione</b>	<b>Ricerca e Sviluppo</b>	<b>Risorse Umane</b>	<b>Acquisti</b>

Tabella 11: Livello della domanda di BDA da parte delle funzioni aziendali

- ✓ *CRM Analytics*: analisi vendite, reclami, customer management, contact center, digital marketing, price optimization;
- ✓ *Customer Experience Analytics*;
- ✓ *e-Commerce*;
- ✓ *Finance & Accounting*: financial performance, strategy management, governance, risk & compliance analytics;
- ✓ *Human Resources*: Workforce Analytics;
- ✓ *IT Operations*;
- ✓ *Production Planning & Sales*: domanda, offerta e pianificazione della produzione;
- ✓ *Security Analytics*: security & legal;
- ✓ *Social & Web Analytics*: behavioral & cohort analysis, sentiment analysis & brand reputation su reti sociali;
- ✓ *Soluzioni verticali per il business*;
- ✓ *Supply Chain Analytics*: procurement, logistica, analisi costi di produzione, fornitori, ordini;
- ✓ *Telecommunication Analytics*: analisi del flusso delle chiamate, assistenza clienti, promozioni;
- ✓ *Top Manager Dashboard Solutions*;
- ✓ *Transportation Analytics*: gestione parco mezzi, gestione carico e distribuzione.

Andando ad osservare come sono distribuite le X nella tabella, emerge che è il Marketing e Vendite a richiedere il maggior numero di soluzioni, pari a 8. Questo risultato non ha disatteso le nostre aspettative considerando l'alto livello di maturità della funzione. In particolare possiamo notare come questa sia l'unica a far domanda di CRM e Customer Experience Analytics che rappresentano una delle applicazioni maggiormente offerte dalle startup. Il Marketing e Vendite è seguito dall'Amministrazione, Finanza e Controllo, dagli Acquisti e dalla Logistica e Produzione che ne richiedono 5 e dalle Risorse Umane che ne domandano 4. La funzione che considera poco queste soluzioni è quella di Ricerca e Sviluppo. Infine tutte sono interessate alle soluzioni verticali per il business e ai top manager dashboard solutions; le differenze risiedono nella richiesta di applicazioni specifiche per la funzione aziendale, come possiamo osservare dalla tabella.

Considerando tutte le soluzioni proposte dalle startup, in modo qualitativo andiamo ora a mappare su una matrice, mostrata nella pagina successiva, il livello dell'offerta, crescente spostandosi verso destra, e quello della domanda, crescente spostandosi verso l'alto, al fine di

individuare il posizionamento di ciascuna funzione aziendale relativamente a queste due dimensioni.

In particolare per la dimensione relativa al *livello dell'offerta* abbiamo considerato le seguenti proxy:

- la numerosità delle soluzioni offerte, calcolata precedentemente per ciascuna funzione aziendale nella tabella 10;
- l'entità dell'offerta di ciascuna soluzione messa a disposizione dalle startup, determinata dalla loro classificazione per ambiti applicativi e riportata nella tabella 10.

Per il *livello della domanda* invece:

- la quantità di soluzioni richieste dalle funzioni aziendali, determinate nella tabella 11;
- lo stato attuale di fruizione delle soluzioni di BDA dalle diverse funzioni, che ritroviamo nella tabella 9 e che abbiamo già commentato, ipotizzando che se una funzione implementa una data soluzione vuol dire che l'ha richiesta.

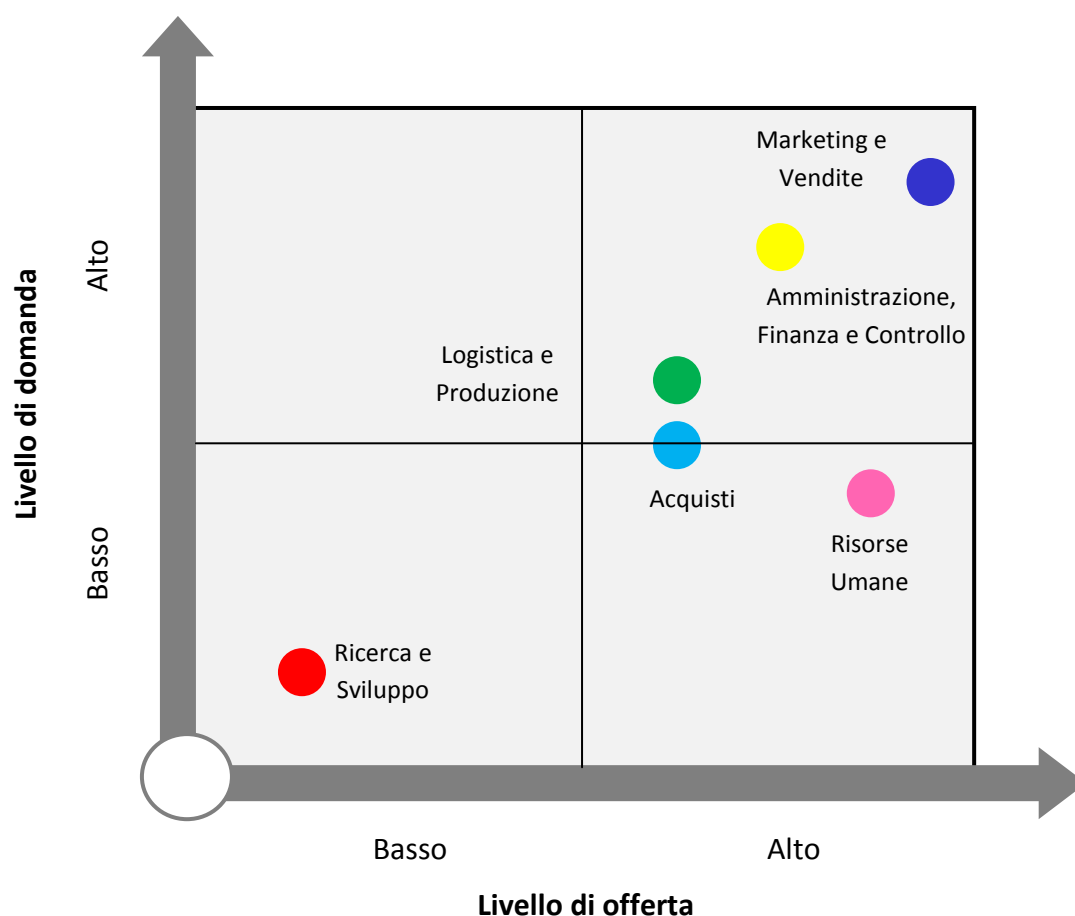


Grafico 22: Posizionamento delle funzioni aziendali rispetto nella matrice livello di domanda/livello di offerta

Il dato che salta subito all'occhio è l'ottima posizione occupata dalla funzione *Marketing e Vendite*, caratterizzata da un alto livello di domanda al quale corrisponde un'offerta adeguata, a testimonianza di come le startup abbiano compreso le esigenze delle aziende e mettano a disposizione in misura maggiore soluzioni rivolte alla relazione con il cliente piuttosto che ai processi interni. Dalla parte opposta della matrice si trova invece la funzione di *Ricerca e Sviluppo*, dove sia la domanda che l'offerta di Big Data Analytics sono scarse: la strada da fare per spostarsi nella zona in alto a destra è ancora molto lunga. L'*Amministrazione, Finanza e Controllo* si trova in buona posizione dove offerta e domanda si equivalgono, ma a livelli inferiori rispetto al Marketing e Vendite, mentre una posizione particolare è quella occupata dalle *Risorse Umane*, dove possiamo osservare un grande traino dell'offerta che propone soluzioni per ottimizzare e migliorare i vari processi, ma che non vede uno stesso riscontro nella domanda. Anche le funzioni di *Logistica e Produzione* e di *Acquisti* sono caratterizzate da un'offerta maggiore della domanda (livello medio), ma si tratta di uno scostamento minore rispetto a quello delle Risorse Umane. Quest'ultimo è inoltre destinato a ridursi considerando che la prima sta iniziando a richiedere delle soluzioni di Big Data Analytics per l'analisi dei costi di produzione, la gestione degli ordini e dei fornitori, il procurement e la gestione del parco mezzi e del carico e la seconda si sta interessando a soluzioni per la gestione della spesa, dei pagamenti, degli ordini e per l'analisi delle performance finanziarie, le quali trovano delle risposte nell'offerta.

Dall'analisi condotta sono due i principali risultati che emergono:

- La funzione di Marketing e Vendite è quella che ha a disposizione l'offerta più ampia e che quindi può trarre i maggiori vantaggi.
- L'offerta in realtà copre le esigenze di tutte le funzioni aziendali, fornendo un supporto adeguato e addirittura, per esempio nel caso delle Risorse Umane, trainando la domanda.

Il posizionamento della maggior parte delle funzioni aziendali all'interno della matrice è coerente con il livello di maturità che abbiamo determinato in precedenza. Chi domanda di più sono le funzioni pronte (Marketing e Vendite) o comunque sulla buona strada (Amministrazione, Finanza e Controllo). Il fatto che l'offerta per le Risorse Umane non veda pieno riscontro nella domanda probabilmente è anche dovuto al suo basso livello di maturità. Non sorprende infine la posizione della Ricerca e Sviluppo.

## 5.5 AMBITI PROGETTUALI DEI BDA DELLE FUNZIONI AZIENDALI MATURE

Entriamo ora nello specifico dei progetti di Big DataAnalytics della funzione aziendale che abbiamo individuato come la più matura, ovvero il *Marketing e Vendite* e descriviamo brevemente quelli delle funzioni con un livello di maturità medio/alto o medio, cioè *l'Amministrazione, Finanza e Controllo*, gli *Acquisti* e la *Logistica e Produzione*

### 5.5.1 MARKETING E VENDITE

Gli ambiti progettuali più diffusi all'interno della funzione sono il Direct Marketing (73%), il Cross-Selling e l'Up Selling (61%) e i Social Analytics (46%). In queste tre aree sono stati compiuti grandi progressi rispetto all'anno scorso dove la presenza in queste tre aree era pari rispettivamente al 53%, 37% e 40% (dalla survey CMO dell'anno scorso); tuttavia mentre per la prima non c'è molto interesse prospettico (32%), per le altre due si prevede ancora una crescita, specialmente per i progetti in ambito Social Media, considerando, come detto più volte, il boom del fenomeno social. Per quanto riguarda quest'ultimo dominio è emerso come le iniziative messe in atto siano per lo più quelle basilari di Social Monitoring e quindi di raccolta delle conversazioni online o di Social Listening che si spinge ad una interpretazione dei dati, ma non prevedono ancora l'integrazione dei dati social con quelli presenti nei sistemi informativi aziendali ed è anche per questo che l'interesse futuro è così alto. La situazione è rimasta pressoché invariata rispetto all'anno scorso, dove la maggioranza del campione intervistato, il 29%, ha infatti dichiarato di aver in atto progetti di Social Listening. Ci sono poi alcuni ambiti emergenti, sebbene ancora poco diffusi, quali il Location-Based Marketing (presenza attuale nell'8% degli indagati e interesse prospettico nel 58%), la Customer Micro-segmentation (presenza attuale nel 35% degli indagati e interesse prospettico nel 60%) e l'analisi dell'efficacia di nuovi prodotti, servizi e business model (presenza attuale nel 22% degli indagati e interesse prospettico nel 42%). Emerge in particolare la volontà di sfruttare in diversi modi tutti i dati, quindi quelli di geolocalizzazione, quelli transazionali, quelli social e i clickstream per fare azioni di advertising mirate, volte all'acquisizione e al mantenimento dei clienti. Se però andiamo a confrontare questi risultati con quelli dell'anno passato, possiamo osservare come le percentuali siano rimaste pressoché invariate e quindi come non siano stati fatti dei passi avanti. Infine altri due progetti in cui lo sfruttamento dei Big Data offre grandi opportunità, ma su cui il Marketing non ha ancora fatto leva, sono il Price Optimization e l'In-Store Analysis (presenza rispettivamente pari al 16% e al 24% del campione e interesse prospettico pari rispettivamente al 40% e al 37%). Le organizzazioni hanno infatti a disposizione, da un lato, un numero crescente di dati sul pricing e sulle vendite, considerando

l'esplosione dei canali di vendita online, per ottimizzare i prezzi e dall'altro un elevato numero di tecnologie quali Wi-Fi e strumenti bluetooth, sistemi dei punti di vendita, carte di pagamento e applicazioni mobile che permettono di raccogliere un enorme quantità di dati dai quali poi gli Analytics possono estrarre insight per fare l'In-store analysis.

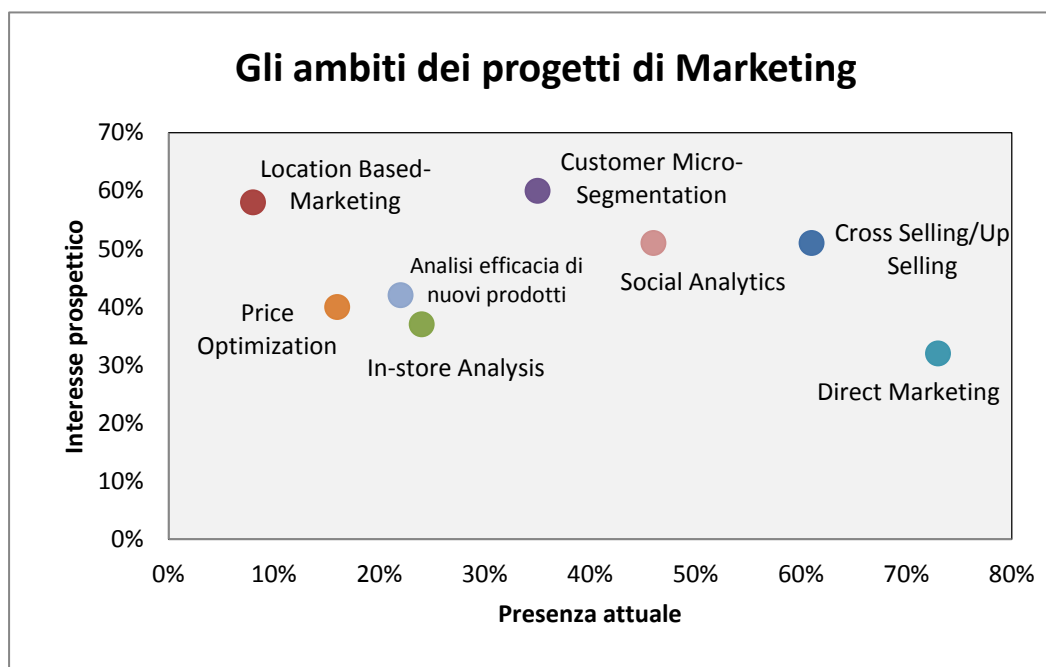


Grafico 23: Posizionamento delle funzioni aziendali rispetto nella matrice livello di domanda/livello di offerta

### 5.5.2 AMMINISTRAZIONE, FINANZA E CONTROLLO

All'interno dell'Amministrazione, Finanza e Controllo per ora viene fatta leva sui Big Data Analytics per pianificare il budget, migliorare l'efficacia delle previsioni e per monitorare i KPI e i livelli di prestazione attraverso dashboard e score-card dinamiche. Al suo interno sono previsti altri progetti futuri, che attualmente si trovano ancora in uno stadio iniziale: si tratta del monitoraggio e delle previsioni continue e dinamiche dei KPI delle divisioni aziendali, dell'analisi del rischio finanziario legato a transazioni con clienti o fornitori e dell'identificazione di frodi esterne.

### 5.5.3 ACQUISTI

La funzione Acquisti adotta soluzioni di Analytics avanzati soprattutto per la gestione amministrativa di ordini e pagamenti, per la tracciatura degli ordini e per l'analisi della spesa. Le

aree in cui l'unità considera di far leva sui Big Data sono lo sviluppo della strategia di acquisto, la negoziazione e la selezione dei fornitori e il monitoraggio delle loro prestazioni.

#### 5.5.4 LOGISTICA E PRODUZIONE

Analizzando la funzione *Logistica e Produzione* emerge che i tool di Big DataAnalytics sono sfruttati per l'analisi dei costi logistici, per la gestione delle attività di magazzino e delle scorte, per l'identificazione di anomalie, per il tracciamento delle spedizioni e per la pianificazione della domanda. C'è ancora un grande interesse verso quest'ultimo ambito e verso l'adozione di soluzioni di Big DataAnalytics per individuare attività anomale, come colli di bottiglia lungo la supply chain per riuscire ad agire tempestivamente ed incrementare ulteriormente l'efficienza, sebbene questo sia un progetto che probabilmente si realizzerà più avanti nel tempo.

### 5.6 CONCLUSIONE

Cerchiamo ora di sottolineare i punti chiave della nostra analisi, facendo riferimento alle domande di ricerca da cui siamo partiti, mettendo quindi in evidenze le aree più avanzate.

Innanzitutto la maggior parte delle organizzazioni ha compreso l'importanza di estrarre insight dai dati e sta iniziando a far leva sui Big Data, ma risulta ancora lontana dall'adottare una strategia di business Data-driven. Inoltre c'è sempre una parte di "scettici", i quali non sono ancora riusciti a cogliere le opportunità da loro offerte.

Considerando le singole funzioni aziendali, è emerso che il Marketing è la funzione più matura nell'ambito dei Big DataAnalytics e con le maggiori prospettive di crescita. Nelle diverse dimensioni indagate, relative allo stato di diffusione dei Big DataAnalytics, della consapevolezza dell'effetto strategico dei Big Data, dei dati utilizzati e delle soluzioni adottate si trova in una condizione migliore rispetto a tutte le altre funzioni, portandola ad occupare una posizione privilegiata. Solo la funzione di Amministrazione, Finanza e Controllo sembra tenere il passo, gli Acquisti e la Logistica e Produzione si trovano in uno stadio intermedio, mentre la Ricerca e Sviluppo e le Risorse Umane sono ancora in una fase arretrata. Considerando quindi la situazione nel complesso possiamo individuare alcuni elementi motivo di soddisfazione e altri meno: se da un lato il lavoro da fare per le funzioni, compreso il Marketing e Vendite, per raggiungere un maggior livello di maturità è ancora lungo, dall'altro possiamo osservare l'impegno di tutte per migliorare.

Le esigenze delle funzioni aziendali sono ben coperte dall'offerta che rappresenta l'area nella quale abbiamo individuato il maggior riscontro positivo. Anche in questo ambito è il Marketing e Vendite a trovarsi nella posizione migliore, considerando l'ingente quantità di soluzioni che gli viene proposta. Le 498 startup censite mettono a disposizione soluzioni che vanno dalle Infrastrutture, agli Analytics fino ad Applicazioni rivolte a uno o più ambiti applicativi e la loro offerta in un mercato così innovativo come quello dei Big Data, è essenziale dato il loro alto livello di innovazione.

Anche l'analisi degli ambiti progettuali delle funzioni più mature mostra una situazione molto promettente, considerando le iniziative previste nel prossimo futuro, che mirano a far leva sempre di più sui Big Data, anche se attualmente si trovano in uno stadio iniziale. Nella funzione di Marketing e Vendite abbiamo infatti individuato alcuni progetti già ampiamente presenti che sfruttano i Big Data, quali quelli di Direct Marketing, di Cross e Up-Selling e di Social Analytics e il grande interesse a sfruttare l'enorme mole di dati a disposizione, tra cui quelli di geolocalizzazione, quelli social e quelli di navigazione sul web per altre iniziative.

Concludiamo presentando i punti di forza e di debolezza e i possibili sviluppi futuri di questo lavoro, che rappresenta un punto di partenza per l'individuazione degli approcci ai Big Data Analytics delle funzioni aziendali delle medie e grandi organizzazioni. Un vantaggio della nostra analisi è rappresentato dal campione significativo delle startup censite che ci hanno permesso di ottenere dei risultati robusti relativamente all'offerta e della significativa quantità di informazioni diverse che abbiamo estratto dalle survey, specialmente da quella rivolta alle funzioni aziendali, che ci ha consentito di trattare numerosi aspetti, dallo stato attuale e previsto di adozione dei sistemi BDA, ai dati analizzati, fino alle soluzioni adottate. Questo ci ha quindi permesso di confrontare le aree di domanda, di offerta e quella relativa al livello di maturità delle funzioni aziendali. Tuttavia il campione non significativo per alcune di esse ha rappresentato un aspetto critico che, in alcuni casi citati, ha portato a risultati dubbi. Attraverso quindi un ampliamento futuro del campione sarà possibile arrivare a dei risultati maggiormente consistenti. Un possibile sviluppo futuro è rappresentato da un'estensione dell'analisi a livello settoriale in modo da identificare il livello di maturità all'interno dei singoli settori per poter comparare le diverse situazioni. Inoltre quando anche le PMI inizieranno ad introdurre soluzioni di BDA al loro interno, sarebbe interessante fare un confronto tra queste e le medie e grandi organizzazioni esaminate. Un ulteriore approfondimento potrà essere fatto quando le



funzioni aziendali si troveranno in uno stadio più avanzato, che prenda in considerazione ulteriori aspetti a quelli trattati, come i benefici quantitativi apportati dalle soluzioni e il contributo dalle nuove figure di Data Scientist e di Chief Data Officer introdotte.

## 6 CASI DI STUDIO

Dopo la parte di analisi, riportiamo ora degli esempi specifici di progetti Big Data implementati da alcune aziende italiane per dar maggior senso ai risultati. A valle della compilazione del questionario, l'Osservatorio ha infatti approfondito circa 70 iniziative italiane, di cui 18 con un'analisi dettagliata tramite interviste telefoniche o de visu con il responsabile del progetto. In queste ultime i tre aspetti su cui ci si è principalmente focalizzati sono:

- *Strategia*: definizione delle esigenze alla base dell'iniziativa, dei processi da supportare, delle tipologie di applicazioni e di servizi introdotti, dei benefici, delle criticità e degli sviluppi futuri e valutazione dell'esistenza di un piano di sviluppo pluriennale.
- *Organizzazione*: analisi delle modalità di introduzione e di implementazione del progetto, del supporto e del coinvolgimento delle funzioni aziendali, delle competenze e dei ruoli emergenti e dell'impatto organizzativo.
- *Tecnologia*: esame delle piattaforme tecnologiche utilizzate e della modalità di implementazione, dell'integrazione dei dati, delle dinamiche di sviluppo e del ruolo dei fornitori.

Gli 8 casi trattati di seguito e altri approfonditi meno dettagliatamente coprono diverse industry, per ognuna delle quali riportiamo brevemente le principali progettualità specifiche, prima di scendere nel dettaglio del singolo caso.

### ASSICURAZIONI

Le Assicurazioni stanno adottando un approccio Big Data orientato al lungo periodo, dedicandosi a progetti quali la gestione delle frodi, la profilazione dei servizi, il pricing delle polizze e la previsione del rischio di abbandono. Il mondo assicurativo contribuisce per il 5% al valore del mercato dei Big Data Analytics, tuttavia presenta il tasso di crescita più alto rispetto agli altri settori, maggiore del 25%. Tra le varie compagnie assicurative italiane, il *Gruppo Generali* sembra aver fatto un enorme passo avanti acquisendo la startup inglese Mydrive Solutions, leader nell'utilizzo degli strumenti di Data Analytics per profilare gli stili di guida in modo da definire prodotti innovativi e smart ritagliati sulle specifiche esigenze della clientela e offrire tariffe vantaggiose per gli assicurati più virtuosi. L'acquisizione ha poi permesso al Gruppo di dotarsi di un centro di eccellenza nell'analisi dei dati (Generali, 2015).

## EDITORIA

Nel campo dell'editoria i principali progetti di Big Data riguardano l'analisi dei profili, la segmentazione dei clienti, il miglioramento dei prodotti editoriali, la personalizzazione dell'offerta, la definizione di azioni di Cross-Selling e Up-Selling. Si tratta di iniziative che rientrano tutte nell'area Marketing. Una nota azienda del settore editoriale, come vedremo dettagliatamente nel caso dell'*Azienda X*, raccoglie i dati dei clienti, da quelli già esistenti sul CRM aziendale a quelli di profilazione e di navigazione, su un'unica piattaforma. Il passo successivo consiste nell'associazione dei dati anagrafici di tutti gli abbonati alla rivista fisica o digitale ai loro comportamenti e gli insight tratti vengono poi trasformati in azioni quali l'invio di offerte personalizzate.

## PUBBLICA AMMINISTRAZIONE

Dalla nostra analisi è emerso come il settore della Pubblica Amministrazione italiano, che contribuisce minimamente al valore del mercato BDA (9%), abbia iniziato a far leva sui Big Data per l'analisi delle frodi e degli errori, per la gestione delle infrastrutture e dei beni demaniali, per il controllo, il governo e l'ottimizzazione degli accessi nelle aree protette delle città e per il rilevamento e il monitoraggio di fenomeni di traffico e di tipo idro-geologico e atmosferico. Un altro interessante ambito progettuale, che approfondiamo nel caso del *Comune di Brescia*, è relativo all'utilizzo di dati di telefonia mobile e di dati dai social e dai web media per fare Marketing territoriale, ovvero per individuare l'effetto che i grandi eventi hanno sulla città al fine di migliorarli, aumentandone l'impatto. La strada da fare per la Pubblica Amministrazione italiana è ancora lunga, visto che è uno tra i settori dove le competenze di Big Data sono meno diffuse e dove le modalità di gestione dei dati sono ancora arretrate.

## RECRUTING ONLINE

Le aziende che si occupano di recruiting online sfruttano i Big Data per fare in modo che le aziende clienti riescano ad avere i migliori candidati e quindi un personale più qualificato. In particolare sistemi di BDA vengono utilizzati per prevedere il tasso di abbandono, le performance, l'assenteismo e il successo dei candidati facendo quindi in modo che le organizzazioni assumano persone con un più lunga permanenza e con una maggiore redditività (caso *Azienda Y*). Altri progetti sono la segmentazione e la profilazione dei candidati, verso cui *Jobrapido* ha dedicato la sua attenzione. L'azienda, come vedremo nell'approfondimento di sotto, si serve dell'indirizzo mail degli utenti e di altri indici che riesce a ricavare per classificarli e profilarli, in modo da offrire il lavoro più adatto a loro. In questo ambito quindi i vantaggi

sono duplici: da un lato le aziende si vedono proporre profili adeguati e dall'altro lato chi è in cerca di un lavoro si vede suggerire delle offerte adeguate alla sua persona e alle sue competenze.

## SETTORE BANCARIO

Dalla nostra analisi è emerso come il settore Bancario sia quello più avanti per quanto riguarda sia le competenze di Big Data sia la strategia adottata per la gestione degli Analytics, esplicita e di lungo periodo. Esso contribuisce per il 29% del mercato e presenta un tasso di crescita elevato (compreso tra il 15 e 25%). I principali progetti di interesse sono l'analisi dell'andamento degli indici azionari, l'analisi sui comportamenti dei clienti, la gestione del rischio creditizio e la gestione delle frodi. Diverse Banche italiane indagate stanno cercando di arricchire le informazioni sui clienti per prevedere il tasso di churn e di tracciare in real time ogni punto e momento di contatto con questi per soddisfare al meglio le loro esigenze. Una nota Banca presente in Italia in cui la grande maggioranza dei contatti agisce tramite i canali online, integra i dati di interazione di 100 mila clienti con il sito web e quello ottimizzato per i dispositivi mobili e l'app mobile al fine di massimizzare l'efficienza del servizio. Nel momento in cui in un cliente ha bisogno di assistenza e chiama il call center, gli strumenti adottati permettono di vedere come questo ha interagito precedentemente con gli altri canali digitali e agire di conseguenza. Quello utilizzato è un Database aperto che permetterà di spostarsi dall'attuale logica reattiva a una proattiva e quindi di fare predictiveAnalytics e programmatic Marketing per esempio, muovendosi con estrema agilità.

## TELCO

Il settore italiano delle Telco si caratterizza per un tasso di crescita nell'investimento in Big DataAnalytics tra i più alti rispetto alle altre industry, compreso tra il 15% e il 25%. Le principali aree in cui tali soluzioni vengono adottate sono l'analisi dei profili e dei consumi, il Marketing geolocalizzato, la customizzazione dell'offerta, il miglioramento dell'esperienza dei clienti, la pianificazione della capacità di rete e la manutenzione, la riduzione degli errori di rete e il supporto servizio ai clienti. Una nota azienda italiana leader nell'InsuranceTelematics si trova in una posizione molto avanzata in quest'area, perseguendo un approccio Data-driven strategico e riuscendo a fornire valore ai suoi clienti. Questa si occupa dell'installazione della black box sensoristica su più di 4 milioni di veicoli circolanti che danno informazioni in real time relativamente al veicolo, ai comportamenti dei guidatori, a dove si trovano e le offrono alle compagnie assicurative. L'ammontare complessivo di dati a disposizione è impressionante: 200

mila record al minuto, mezzo miliardi di dati al giorno. L'azienda si è dotata circa due anni fa di persone e infrastrutture adeguate per fare predictiveAnalytics e quindi per lavorare in tempo reale enormi quantità di dati eterogenei, in modo che le assicurazioni clienti sappiano cosa succederà nel prossimo futuro e quindi quali sono e quali saranno i guidatori più virtuosi, quali i più rischiosi e in quali circostanze. Questo rappresenta un fattore competitivo molto importante per le compagnie assicurative che consente loro di fare pricing e Marketing migliori, traendone grande vantaggio. L'azienda sta inoltre adottando tecniche e tecnologie BDA per ottimizzare i suoi prodotti e processi interni, ricavando profitto da questi servizi a valore aggiunto: dotare la sensoristica di brain in modo che possa modulare, in funzione dell'ambito in cui agisce, in maniera predittiva grazie ai Data insight è l'obiettivo.

## TURISMO

Anche il mondo del turismo online italiano, che può trarre un enorme vantaggio dai Big Data, ha iniziato a far leva su questi. La profilazione dei consumatori, la personalizzazione dell'offerta, il miglioramento della customerexperience e l'ottimizzazione del rapporto con i search e metasearch sono le principali iniziative. Di seguito riportiamo i casi di *Lastminute.com* e *Venere.com*, aziende appartenenti a questo settore, che stanno implementando questi progetti: le informazioni relative alle preferenze e ai comportamenti degli utenti e alle interazioni tra i due siti e i motori di ricerca vengono sfruttate per migliorare l'esperienza dei clienti e per ottimizzare le prestazioni.

## UTILITY

Nel settore delle Utility, che tra gli altri è uno di quelli che investe meno in Analytics, i principali progetti riguardanti i Big Data sono l'analisi dei consumi, la pianificazione delle capacità e la manutenzione predittiva della rete, la valutazione del rischio, il Marketing geolocalizzato e l'analisi del tasso di churn. Quest'ultimo è un ambito in cui si è impegnata *Eni*, la quale, come descritto meglio nel caso riportato di seguito, ha sviluppato modelli predittivi per le commodity offerte per individuare il comportamento del singolo consumatore in modo da determinare il suo tasso di abbandono, perseguendo una logica cliente-centrica. L'azienda fa inoltre leva su dati a disposizione per il campaign management e quindi per implementare su tutti i canali azioni adeguate e per passare le informazioni corrette a clienti che chiamano inbound perché hanno dei problemi da risolvere.

Presentiamo ora i casi approfonditi.

## 6.1 ADVICE GROUP

Advice Group è l'unica agenzia di Progress Marketing italiana che sfrutta le informazioni derivanti dalle attività promozionali per gestire programmi complessi di behavioral e real time loyalty. Si occupa di creare programmi di relazione digitale che hanno l'obiettivo di migliorare le performance di sell-in e sell-out dei propri clienti. Monitorando in corsa l'andamento dei progetti è possibile mappare il comportamento degli utenti nel corso delle attività per poter offrire opportunità personalizzate nel momento più favorevole alla loro conversione a favore del ROI. Aiutano quindi le aziende ad arricchire in modo dinamico il profilo dei propri utenti e clienti, introducendo il concetto di monetizzazione dell'informazione relazionale, attraverso la cessione di opportunità personalizzate in tempo reale. Offrono soluzioni rivolte all'area consumer, trade, Retail e MICE. La sede centrale è a Torino e hanno deciso di allargarsi ed espandere il loro mercato anche in Sud America partendo dal Perù con una sede a LIMA, paese in cui si prevede che il PIL supererà la media mondiale entro il 2020, e dove il mercato digitale è molto giovane: il 62% degli utenti connessi è shopper tra i 25 e i 54 anni. Nel 2014, grazie ad un progetto realizzato per un cliente, hanno vinto il premio come "Miglior Progetto B2B nel mercato EMEA" ai Loyalty Awards di Londra.

La necessità di confrontarsi con il mondo dei Big Data nasce nel momento in cui l'azienda riscontra un gap nel mercato digitale, rappresentato sia dall'assenza di una strategia integrata nell'utilizzo dei diversi touchpoints fisici e digitali, sia dall'idea che gli investimenti in digital siano solo legati all'entertainment e non portino un effettivo ritorno in termini di vendite. L'idea è quindi offrire ai propri clienti soluzioni di Progress Marketing che permettano loro di sfruttare al meglio le grandi moli di dati raccolti, per conoscere maggiormente i consumatori e riuscire ad instaurare delle relazioni sempre più durevoli e customizzate, incrementando il sell-out.

Un progetto particolarmente interessante sviluppato nell'ambito Big Data è WEKIT, WeKeep In Touch. Si tratta di una piattaforma di behavioralloyalty implementata e sviluppata anche con il supporto dell'azienda Consoft, realtà dedicata al mercato B2B.

Wekit è un prodotto che nasce per mettere ordine nella confusione digitale delle aziende, in quanto è in grado di connettere fra loro tutte le attività digitali e promozionali realizzate dal brand in un unico repository. Tutti i dati raccolti vengono analizzati, elaborati e riclassificati e resituiscono *Key Performance Indicator* comportamentali degli utenti coinvolti, informazioni utili per attività di Marketing e CRM successive.

Wekit è una Data Management Platform che si distingue da altri progetti di DMP per il suo orientamento al settore Promotion. Infatti, oltre ai dati provenienti da touchpoints fisici e digitali, dall'evento ai social network, è in grado di raccogliere anche le informazioni in ambito promozionale, da semplici contest a complessi loyaltyprogram. I dati vengono immagazzinati e normalizzati in modo da renderli confrontabili anche con le informazioni di vendita fornite dai clienti. Wekit restituisce così cluster comportamentali degli utenti, che tengono conto sia del comportamento di acquisto che della predisposizione di ogni utente all'utilizzo di un device, della sua reattività verso stimoli proposti dal brand e della sua potenzialità virale. Queste informazioni rendono i progetti digitali più rispondenti perché la clusterizzazione permette di massimizzare il tasso di risposta alle missioni erogate ai consumatori. La piattaforma è in grado di gestire contest, loyaltyprogram, incentive trade, e-commerce, community e azioni one-to-one personalizzate e mirate. Aiuta a misurare con maggior precisione redemption, Roi e Ror trasformando l'attuale visione statica del CustomerRelationship Management in una maggiormente dinamica di CustomerRelationshipStrategy.

L'obiettivo del progetto, nonché fattore critico di successo, è quello di realizzare progetti di Progress Marketing. Con questo termine si intende il collegamento tra le attività di digital Marketing e il potenziamento del sell out. In questo modo il focus non è concentrato solamente sulla brand awareness ma anche sul CrossSelling fra prodotti. Integrando ogni azione digital con il sell-out si passa dall'ingaggio alla fidelizzazione, trasformando semplici follower in clienti reali.

La collaborazione con Consoft ha riguardato l'attività consulenziale per la scelta degli strumenti di analisi e rappresentazione dei dati in ambito Big Data, e per il passaggio a Database e attività NoSql.

Le criticità affrontate nella realizzazione della piattaforma sono state maggiormente di carattere tecnologico, poiché è stato necessario trovare la tecnologia giusta che fosse contemporaneamente consoliData e solida. In un mercato come quello tecnologico che varia molto e che presenta in Italia poche esperienze e case history con cui confrontarsi, è stato difficile trovare gli strumenti che potessero essere allo stesso tempo solidi e consolidati, è stato quindi importante l'aiuto di Consoft nella selezione della tecnologia. Per quanto riguarda la tematica della privacy e della sicurezza, questa non è stata vista come una problematica, ma bensì come una sfida da affrontare.

Nel futuro Advice Group vuole garantire che le tecnologie alla base di Wekit siano altamente scalabili, in modo che i moduli della piattaforma siano a loro volta dimensionabili a grandi numeri e Cross countries, garantendo allo stesso tempo la sicurezza dei dati, elevate performance e internazionalizzazione del knowhow dell'azienda Torinese.

## 6.2 AZIENDA X

L'azienda X è tra le principali società europee nel settore editoriale, la cui missione consiste nel favorire la diffusione della cultura e delle idee con una produzione che tocca ogni genere e raggiunge tutti i lettori. Le sue attività sono articolate in diverse business unit tra cui quella Digital trasversale alle altre attività, il cui obiettivo è rafforzare la presenza dell'azienda nel mercato digitale e accelerare i processi di innovazione.

L'esigenza di sfruttare i Big Data nasce nell'unità Digital: essendo il fatturato digitale della società più basso rispetto a quello degli altri player europei, questa ha pensato di far leva sui dati dei clienti per incrementare la revenue al fine di raggiungere e superare i diretti competitor. Si tratta di un'ottima soluzione considerando la customer base di 25 milioni di clienti e quindi l'enorme mole di informazioni a disposizione.

A tal fine è stato implementato un progetto triennale di Big Data all'interno di una più ampia strategia Data-driven che, dopo una prima fase di sperimentazione, oggi si trova nella fase iniziale di set-up ed ha già dato risultati positivi. L'obiettivo è quello di elaborare il grande deposito di dati e trasformarlo in valore maggiore per il business. È stata quindi adottata un'unica piattaforma di gestione dei dati (DMP) fornita dall'azienda Y per tutte le aziende del gruppo multimediale a cui appartiene X, che permette di gestire l'insieme dei dati online e offline e di mettere a fattore comune le richieste espresse dal Gruppo.

La scelta di questa soluzione è legata a due diversi aspetti: in primo luogo la piattaforma unica consente di fare Datawarehousing in modo tradizionale su grosse moli di dati con ottime prestazioni, superando i problemi relativi all'alto costo di memorizzazione e alle performance povere dei sistemi attuali, su cui i dati vengono cancellati e sintetizzati in maniera eccessiva. In secondo luogo si tratta di una soluzione compatibile ai tool di BI e business analysis anche tradizionali e non solo Big Data e quindi più solida rispetto ad altre. I proof of concept realizzati sono andati a buon fine: è stata fatta una sperimentazione in cui è stato dedicato a questa soluzione un singolo nodo di una piccola macchina virtuale e su questo sono stati fatti girare i report del Business Object che insistevano sul Database utilizzato. Il risultato ottenuto è stato un miglioramento delle performance di 5 volte che ha portato alla scelta di questa DMP.



I dati raccolti su questa piattaforma esterna sono sia strutturati che destrutturati: si tratta di dati di profilazione e di navigazione dei clienti i quali accedono al sito in modalità differenti, quindi dati anonimi sul comportamento digitale che arricchiscono il patrimonio informativo già esistente sul CRM aziendale. La DMP consente di associare i “freddi” dati anagrafici di tutti gli abbonati alla rivista fisica o digitale ai loro comportamenti; questi vengono poi trasformati in azioni rivolte a loro che possono essere pianificate, erogate e misurate in real time. Il fatto di avere una piattaforma unica moltiplica l’audience ed è molto importante, tuttavia non è sufficiente per trarre tutti i vantaggi dalle informazioni a disposizione e per questo motivo vengono create anche altre piattaforme aziendali per la gestione dei dati personali dei clienti sui quali costruire progetti quali gli smartengines, ovvero i motori di raccomandazione. Inoltre i dati relativi alle caratteristiche e alla navigazione degli utenti, per esempio, possono essere utilizzati per creare dei segmenti e quindi essere maggiormente valorizzati sulle altre piattaforme di programmatic advertising.

Il progetto è stato seguito da un Data scientist, inserito in azienda nel 2014, il quale ha stimolato la fase sperimentale e si è occupato della realizzazione dei prototipi.

L’implementazione di questa soluzione Big Data permette di raggiungere una serie di benefici. Un primo grosso vantaggio riguarda l’advertising in quanto l’utilizzo dei dati permette alla società di vendere al meglio la sua pubblicità, soprattutto sulle piattaforme programmatiche, ottenendo un incremento del 20-30% sul cost per mille (CPM), ovvero la stima del costo di 1000 visualizzazioni di un messaggio pubblicitario online. La seconda potenziale linea di revenue è quella ottenuta dalla vendita diretta dei dati da parte di società specializzate in direct Marketing che operano all’interno dell’azienda. Un terzo beneficio derivante dall’arricchimento della customer base è rappresentato dalla possibilità di conoscere il cliente a 360 gradi, il che permette di migliorare i prodotti editoriali, di personalizzare le offerte e di fare azioni di up-Selling e Cross-Selling, aumentando le opportunità di contatto diretto con il cliente. Quest’ultimo aspetto impatta positivamente sulla customer experience e quindi sulla soddisfazione dei consumer, che rappresenta un fattore critico di successo dell’iniziativa.

La base degli ottimi risultati del progetto è comunque rappresentata dall’adozione della piattaforma unificata per tutte le aziende del Gruppo, caratterizzata da ottime prestazioni e compatibile con tutti gli altri strumenti utilizzati dalla società.

Nel corso della realizzazione del progetto sono state affrontate una serie di criticità: le principali sono quelle culturali, dovute a lacune nel know-how all’interno dell’organizzazione

ovvero alla mancanza di figure quali Data scientist e professionisti del dato ed è anche per questa ragione che la durata della fase sperimentale è stata così lunga. Inoltre all'interno dell'azienda l'approccio alla cultura è molto legato al singolo business: il concetto di utilizzo del dato come possibile leva per fare altro business e per ottenere dei benefici e come valore Cross, patrimonio di tutta l'azienda e non delle singole unità organizzative è ancora poco sviluppato. Il Gruppo ha poi dovuto superare ostacoli relativi alla regolamentazione e alla privacy, mentre da un punto di vista tecnologico non sono stati riscontrati particolari problemi.

Per il futuro la Società intende innanzitutto portare a termine il progetto Big Data implementato, dal quale si aspetta di estrarre il maggior valore nel 2016. Altro obiettivo, legato ai problemi di governance trattati, è l'inserimento di Data scientist e la creazione di un centro di competenze interne su queste tematiche, che permettano di convertire il personale già presente in specialisti dei dati, adottando quindi un approccio misto, sia esterno che interno.

### **6.3 AZIENDA Y**

L'azienda Y è il leader nelle applicazioni cloud per il talent management: aiuta le organizzazioni a reclutare, formare, gestire e connettere i dipendenti, aumentandone la produttività. Lavora con centinaia tra le più grandi imprese in tutto il mondo e con migliaia di piccole aziende al fine di coinvolgere e potenziare la loro forza lavoro e questa è proprio la sua mission.

Oggi le soluzioni sono adottate da 20 milioni di utenti in 191 Paesi e in 42 lingue differenti.

Dal 1999, anno della sua fondazione, l'azienda è stata un innovatrice nelle tecnologie cloud, unico modo attraverso il quale offre i suoi prodotti. I clienti dotati di un puro Software-as-a-Service, hanno sempre l'ultima versione e riescono quindi a raggiungere benefici quali la riduzione dei costi IT e di manutenzione, l'integrazione con altre piattaforme tecnologiche e l'accesso al software dell'organizzazione in qualsiasi momento, in qualsiasi posto e tramite qualsiasi device.

Considerato il mondo in continua evoluzione, l'azienda nel luglio 2014 ha deciso di introdurre una soluzione di Big Data Analytics per offrire qualcosa di innovativo al cliente, che potesse portare lui un valore aggiunto, nonostante la gestione delle risorse umane sia un business tradizionale. Il progetto è stato possibile grazie all'enorme mole di dati a disposizione: ogni giorno infatti 20 milioni di utenti visitano la piattaforma. Le informazioni, tutte presenti nella banca dati, sono relative a:

- Utente: posizione, istruzione, carriera, skill, lingua, ecc...
- Organizzazione: settore, business unit, regione, localizzazione, ecc...
- Collaborazione: team, discussioni, feedback, ecc...
- Talento: performance, skill, obiettivi, assessment, ecc...
- Dati esterni

L'obiettivo è stato mettere a fattore comune questo patrimonio di dati con il talent management, passando quindi da analisi statistiche ad analisi predittive.

A tale scopo, sono stati creati gli Analytics Y: machine learning che applica sofisticati modelli predittivi e complessi algoritmi psicologici e attitudinali, basati su 6 anni di esperienza del laboratorio R&D della sede principale, ai grandi insiemi di dati per identificare modelli e fare previsioni nei diversi ambiti di gestione delle RU, in particolare nei tre seguenti:

1. Recruiting: viene previsto il tasso di abbandono, le performance e l'assenteismo dei candidati, basandosi sui comportamenti e sulle valutazioni psicologiche.
2. Learning: vengono forniti suggerimenti per la formazione basandosi sulle performance, sulle relazioni e sulle skill.
3. Performance: viene previsto il successo dei dipendenti nei ruoli futuri, i potenziali successori nelle diverse posizioni e il tasso di uscita, basandosi sui dati attuali della forza lavoro.

Uno dei principali vantaggi della costruzione di un sistema unificato di talent management è avere un'unica fonte di dati relativi a tutti i processi di gestione delle RU alla quale i manager possono attingere.

Report e dashboard consegnano i dati chiave ai clienti attraverso la suite Y, consentendo loro di avere il dato giusto al momento giusto e di migliorare il processo decisionale grazie all'accesso alle ultime informazioni, possibile dato che i report sono aggiornati in real time.

Più in dettaglio i predictiveAnalytics permettono di:

- Scoprire fattori predittivi di non conformità alle leggi e ricevere dei suggerimenti di cambiamento delle policy per ridurre il rischio delle organizzazioni di incorrere in pene legali e per evitare i tempi di inattività dei dipendenti.
- Identificare percorsi di carriera di successo per i dipendenti, aumentando quindi il loro impegno e le loro prestazioni.

- Determinare quali fattori predicono alte prestazioni e quale struttura di ricompensa le migliora.
- Individuare fattori predittivi del rischio di abbandono dei dipendenti, permettendo quindi di trattenere quelli con le migliori performance e di ridurre il tasso di uscita, che comporta alti costi di sostituzione, perdita di conoscenza e ritardo dei progetti.

Le aziende clienti traggono quindi una serie di benefici dagli insight estratti. Il primo consiste nel gestire proattivamente la forza lavoro, in quanto essi forniscono strumenti che suggeriscono anticipatamente delle modifiche da implementare e visualizzano il loro impatto. Gli insight inoltre si trasformano in azioni da mettere in atto per tenere i dipendenti all'interno dell'azienda il più a lungo possibile e per incrementare le loro prestazioni. Infine essi rendono i business leader molto più strategici: sfruttando insight in tempo reale, questi possono migliorare le prestazioni delle varie unità di business e i KPI di centinaia di dipendenti di vari gruppi e posizioni.

Le soluzioni offerte finora sono trasversali, il livello di personalizzazione è per settore e geografia e non per singola azienda. Il basso livello di customizzazione richiesto è dovuto alla novità delle soluzioni e molto probabilmente aumenterà con il tempo.

Il fattore critico di successo di questo progetto di Big Data è rappresentato dal valore concreto delle soluzioni Y, che comportano un positivo ritorno sull'investimento delle aziende clienti, quali un aumento del tasso di retention. Costi ridotti, migliori prestazioni, maggiore profittabilità sono vantaggi che spingono le aziende a rivolgersi alla società Y. Altro elemento di successo è l'innovatività della soluzione, considerato un settore tradizionale come quello di gestione delle risorse umane, apparentemente lontano dai Big Data.

La fortuna dell'organizzazione è stata anche quella di avere a disposizione un grande volume di dati strutturati in quanto i 20 milioni di utenti erano già contenuti in un Database strutturato; dati che sono stati poi destrutturati per rendere il motore più efficiente e gli Analytics più veloci, visto che la maggior parte delle analisi predittive viene realizzata in real time.

La società può oggi attingere immediatamente a dati altamente granulari e accurati relativi a preziosi talenti, raccolti per più di 15 anni e ad estrarli producendo delle raccomandazioni ai clienti soprattutto nell'area della formazione e dello sviluppo. Questo cambierà lo sviluppo professionale, accelerando l'apprendimento delle persone, le quali sceglieranno dove esplorare nuove conoscenze. Tutti i vantaggi apportati dagli Analytics Y, hanno aumentato la customersatisfaction, comportando una aumento del fatturato e del numero di clienti.

Gli obiettivi futuri sono utilizzare anche dati da fonti esterne come Facebook, Twitter e LinkedIn; tuttavia ci sono due elementi critici da considerare per perseguire questo scopo. Il primo è la privacy, in quanto i clienti sono molto sensibili ai loro dati personali e non vogliono che questi vengano utilizzati dalle aziende soprattutto perché non conoscono esattamente l'uso che ne verrà fatto. Il secondo elemento è rappresentato dal fatto che le soluzioni offerte sono in cloud: 20 milioni di utenti girano su un'unica istanza e se va in crash il sistema dopo che questo è stato aperto, l'azienda ha grossi problemi dato che garantisce SLA al 97% 24 x 7. Se i server non funzionano più la colpa viene attribuita alla società Y e non a Facebook o a Twitter. È fondamentale quindi mantenere il livello di servizio.

L'azienda intende poi sfruttare progressi futuri nei semanticAnalytics al fine di rilevare automaticamente profili online come LinkedIn e curriculum per identificare parole, indici di alte performance. Un altro obiettivo, considerato l'enorme volume di dati generati dai dipendenti ogni giorno, è quello di trasformare il processo di revisione delle performance da un evento annuale a un processo continuo di interazione con i dipendenti, che ricevono dei feedback basati sui dati.

#### **6.4 COMUNE DI BRESCIA**

L'attenzione per la *smart city* ha assunto negli ultimi tempi grande importanza anche in Italia. Dal 2013 il Comune di Brescia ha intrapreso varie iniziative collegate all'approccio *smart city* ed è in quest'ottica che si inserisce il progetto di valutazione, anche attraverso l'utilizzo di *Big Data* telefonici, dell'impatto mediatico, sociale ed economico dei grandi eventi cittadini (e tra questi la Mille miglia storica in primo luogo).

Il progetto, che ha valenza di Marketing urbano, nasce con l'obiettivo di valutare l'impatto di grandi eventi promossi in città e ottenere indicazioni utili a migliorare la qualità di questi eventi e il rapporto con le associazioni di categoria.

L'evento su cui si è concentrato l'interesse è la Mille Miglia storica, manifestazione automobilistica storica italiana che annualmente nel mese di maggio prevede la percorrenza di mille miglia (Brescia –Roma andata e ritorno) su automobili d'epoca prodotte non oltre il 1957.

Il progetto ha preso avvio nel 2013, anno in cui oltre alla Mille miglia storica Brescia visse l'arrivo dell'ultima tappa del Giro d'Italia, e prosegue da allora. Questa prima esperienza è stata sviluppata grazie all'essenziale collaborazione tra l'Ufficio di Staff Statistica del Comune di Brescia e DSM StatLab - Laboratorio di Statistica "Dati Metodi e Sistemi dell'Università degli

studi di Brescia, che nel caso della Mille Miglia hanno contribuito alla raccolta e all'analisi dei dati di telefonia mobile, integrati con i dati provenienti da rilevazioni tramite questionario, con l'obiettivo di fornire un profilo degli spettatori della manifestazione, dei cittadini e dei negozianti coinvolti in modo diretto o indiretto. Attraverso l'utilizzo di dati relativi alla presenza di utenti agganciati alla rete di telefonia mobile, ovvero dati geo-referenziali e anonimi, rilevati a intervalli regolari di 30 minuti (per gli eventi del 2013) e addirittura 15 minuti (per gli eventi dal 2014 in poi), è stato possibile monitorare un'area che comprende il centro storico di Brescia ottenendo una stima della densità di utenti di telefonia mobile presenti. Dal punto di vista statistico, la tecnica della gap analysis applicata a serie storiche ha permesso di estrarre il profilo temporale (ogni 30 minuti) della distribuzione delle presenze sul territorio comunale attribuibili alla manifestazione Mille Miglia o Giro d'Italia in termini di scostamento delle presenze dalla giornata tipo. Vi sono state alcune criticità durante lo sviluppo del progetto legate a problemi di raccolta, conservazione e trattamento delle informazioni, soprattutto all'aumentare della quantità di dati a disposizione.

La prima sperimentazione relativa alla Mille Miglia 2013, oltre ad esser stata una sorta di esercizio per capire e risolvere problemi tecnici e metodologici connessi all'utilizzo dei dati, è stata utilizzata per la programmazione della Mille Miglia 2014. Le scarse presenze rilevate la sera di sabato, giorno dell'arrivo, hanno fornito il supporto informativo alla decisione del Comune di spostare la premiazione della gara alla domenica, così da favorire l'organizzazione di eventi collaterali, nella convinzione di poter aumentare la partecipazione e valorizzare meglio la città dal punto di vista turistico anche durante le fasi conclusive della storica gara. Una seconda sperimentazione è stata svolta durante l'altro grande evento sociale svoltosi a Brescia la settimana successiva all'arrivo della Mille Miglia, cioè l'ultima tappa del Giro d'Italia 2013.

Grazie a un progetto di questo tipo è possibile svolgere analisi mirate a livello sia territoriale sia temporale, considerando rispettivamente particolari aree geografiche e fasce orarie: nel caso di eventi particolari, quali la Notte Bianca o la Festa della Musica, queste informazioni possono essere utili per definire percorsi e strategie di Marketing territoriale, soprattutto se completate con rilevazioni statistiche sul campo.

Il progetto è proseguito nel 2015 allargando l'orizzonte di analisi ed è in questo contesto che si inserisce la collaborazione con Cineca, che ha effettuato la raccolta e l'analisi dell'impatto della Mille miglia sui Social. Cineca, in collaborazione con l'Osservatorio di Pavia – Media Research, ha condotto l'analisi in particolare sui social media Twitter e Instagram, oltre che sui web media. Sui social networks, una volta individuati dei criteri di estrazione dei post, sono state effettuate

delle analisi di text mining, nel dettaglio Cluster analysis, Sentimentanalysis e Network analysis. Nella cluster analysis, sono stati raggruppati i post in base alle parole in essi contenuti. Una volta definiti i cluster è stato necessario un lavoro interpretativo per collegarli a aspetti, eventi o problemi della realtà rappresentata: in questo caso, la corsa e il suo intorno, costituito da geografia, sponsor, passione degli spettatori. Nella sentimentanalysis è stata attribuita una polarizzazione ai post presi in considerazione. A ogni cluster, è stato collegato un indice di sentiment, ottenuto attraverso la ponderazione tra lemmi considerati “positivi” e lemmi considerati “negativi” all’interno di ciascun cluster. Per quanto riguarda la Network Analysis, l’obiettivo era comprendere, attraverso l’analisi del contenuto dei messaggi, le relazioni tra gli utenti per poter quindi visualizzare queste relazioni in modo grafico. Infine è stata eseguita un’analisi sulla geolocalizzazione dei posts di Instagram per individuare le zone di maggiore affluenza.

Analisi dei dati di questo tipo, che comprendono sia i dati relativi all’utilizzo di una rete di telefonia mobile sia quelli social, sono molto utili anche in altri ambiti nei quali l’amministrazione locale è impegnata, per sviluppare modelli spazio-temporali dedicati a studi di mobilità urbana e di impatto ambientale. Le premesse sono quindi molto positive, visto che le tecnologie digitali offrono oggi e offriranno sempre di più in futuro alla statistica moderna la possibilità di sfruttare le sue grandi potenzialità metodologiche per trasformare dati in informazioni, così da favorire decisioni più consapevoli, quindi più efficienti e più efficaci.

## 6.5 ENI

Eni, impresa integrata che opera in tutta la filiera dell’energia con più di 84.000 dipendenti in 83 Paesi nel mondo, rappresenta una delle maggiori multinazionali energetiche globali e la prima industria italiana. E’ attiva nell’esplorazione, sviluppo ed estrazione di olio e gas naturale. Per mezzo di raffinerie e impianti chimici l’azienda processa greggi e cariche petrolifere per la produzione di carburanti, lubrificanti e prodotti chimici venduti all’ingrosso o tramite distributori. Opera nel settore dell’ingegneria e costruzione di impianti e infrastrutture onshore e offshore, concentrandosi sulla realizzazione di progetti tecnologicamente avanzati. Inoltre Eni offre il servizio di car sharing noto come “Enjoy” nelle città di Milano, Roma, Firenze e Torino.

Una delle principali divisioni dell’azienda è la business unit Gas &Power che si è trovata per prima ad affrontare il tema Big Data. La spinta a confrontarsi con la tematica è nata principalmente da due fattori. Da una parte la presenza di un’enorme mole di dati a disposizione dell’impresa grazie agli 8 milioni di clienti, 10 milioni se si considerano anche le

consociate estere, che realizzano giornalmente interazioni e transazioni da cui è possibile ricavare informazioni. Dall'altra l'apertura del mercato dell'energia, che ha indotto Eni ad aumentare gli sforzi in questa direzione per mantenersi competitiva sul mercato. È emersa quindi nel tempo l'importanza di studiare al meglio la propria customer base per poter implementare, nei tempi opportuni, azioni adeguate per ottenere vantaggi rispetto ai nuovi competitor e confermarsi leader.

Per soddisfare queste esigenze l'azienda ha deciso di realizzare un progetto di Big Data. Il percorso seguito ha previsto l'integrazione dei dati provenienti da diversi sistemi informativi aziendali e la loro successiva rielaborazione per ottenere informazioni profilate per singolo cliente, costruendo una visione maggiormente cliente-centrica, e per sviluppare modelli predittivi che potessero determinare il comportamento del consumatore e il relativo tasso di churn. Per raggiungere questo risultato sono stati seguiti diversi step. Inizialmente sono stati sviluppati dei modelli predittivi per la parte luce, quindi quelli per la parte gas e poi per entrambe le commodity, considerando sia i clienti storici che quelli nuovi, ovvero quelli acquisiti sul mercato e maggiormente inclini al cambiamento. Si è arrivati quindi a valutare il cliente nella sua completezza superando la lettura dei dati in chiave di singolo prodotto. Infine l'ultima fase ha riguardato la comprensione del valore ottenibile dai dati così raccolti ed elaborati, potendo ottenere risultati più precisi relativi ai consumi veri e propri. E' stato inoltre realizzato un sistema di visualAnalytics costruito sull'intera base di dati che ha permesso la divulgazione e la condivisione ad ogni livello aziendale delle analisi descrittive sulla customer base, anche mediante l'utilizzo di dispositivi mobili.

I dati utilizzati in questi modelli predittivi sono moltissimi, basta considerare le milioni di variazioni a livello settimanale delle informazioni relative ai 10 milioni di clienti. Si tratta per lo più di dati strutturati legati al comportamento d'acquisto, di contatto e al comportamento digitale del consumatore, come dati di navigazione e di utilizzo delle applicazioni mobile di Eni. Le fonti da cui provengono tali dati sono sia interne che esterne, tra quest'ultime l'Authority, che fornisce informazioni relative al mercato regolamentato. E' stato sviluppato da Eni, in collaborazione con delle aziende partner, un sistema di analisi per confrontare i dati relativi ai propri clienti con i dati esterni riguardanti il potenziale di mercato, sia a livello di famiglie sia a livello di consumi. Questo permette di monitorare in maniera costante la market share e le performance di Eni nelle diverse aree di business, avendo a disposizione un utile strumento per conoscere dove è necessario intervenire e che rappresenta una base di partenza per comprendere anche come intervenire.



Questo progetto Big Data è stato gestito, e continua ad essere gestito, congiuntamente dalla funzione Marketing e dalla funzione ICT. Per riuscire a sfruttare ed estrarre valore dai dati la collaborazione tra queste due unità è fondamentale, sottolineando la necessità sia di competenze tecniche che di competenze commerciali. Un primo fattore alla base del successo dell'iniziativa è proprio l'integrazione tra questi due mondi. Un secondo è legato al cambiamento culturale all'interno dell'azienda: la costruzione del sistema di visualAnalytics ha permesso di divulgare l'importanza del dato e di rendere disponibile le informazioni a tutti i livelli cambiando, in positivo, il modo di lavorare delle persone, con conseguenti vantaggi. In alcuni casi sono state consolidate e ampliate le competenze delle persone e sono nate figure di Data scientist che raccolgono conoscenze di statistica, matematica, informatica, di programmazione e di business, capaci di creare in autonomia modelli predittivi e in grado di estrarre da essi le informazioni chiave. Il vero elemento differenziante è stato riuscire a raggiungere i primi benefici, anche in termini di introduzione di valore in azienda, in tempi rapidi e in un mondo, come quello delle utility, particolarmente competitivo: i primi risultati si sono visti entro nove mesi dal lancio. Eni, grazie a questo progetto, continua a contenere il tasso di abbandono e quindi a proteggere la propria base clienti mantenendo la posizione di leader.

Le maggiori criticità nella realizzazione del progetto sono state riscontrate nella fase iniziale di strutturazione del Data warehouse con logica cliente-centrica. Da una parte dovute ad un certo scetticismo sulla possibilità di ottenere dall'investimento dei reali benefici interni, dall'altra dovute a ostacoli di tipo organizzativo legati all'alto numero di interlocutori coinvolti. Il fatto però che Eni sia stata in grado di mettere in campo dei vantaggi rapidamente, ha dato credibilità al progetto, superando lo scetticismo e gli ostacoli.

Il passo successivo per l'azienda sarà cercare di sfruttare i Big Data aziendali in una logica integrata considerando sempre di più il mondo Eni nella sua completezza. L'obiettivo è quello di ragionare sul cliente nel suo complesso per portare sempre più beneficio ai clienti fidelizzati al brand, ovvero quelli che usufruiscono contemporaneamente dei numerosi servizi offerti dall'azienda. Le prossime sfide che l'azienda vuole affrontare saranno la sempre maggiore integrazione con fonti dati esterne e la gestione di dati sempre meno strutturati come quelli provenienti dai social media con l'obiettivo di perfezionare sempre più la customerjourney.

## 6.6 JOB RAPIDO

Jobrapido, azienda fondata nel 2006 a Milano, è il leader mondiale tra i motori di ricerca dedicati al lavoro la cui vision è di rivoluzionare il modo in cui i jobseeker di tutto il mondo ottengono una nuova occupazione mediante l'utilizzo del web. L'azienda italiana acquisisce gli annunci di ricerca lavoro pubblicati separatamente su migliaia di siti web, come le pagine dedicate delle singole società private (alla voce "Lavora con noi" o similare), o le bacheche di lavoro delle staffing agencies o delle Jobboards, restituendo ai jobseeker le offerte, classificate e ordinate per rilevanza, in base alla ricerca effettuata. L'azienda opera globalmente in 58 paesi e ha oltre 60 milioni di utenti registrati; nel corso del 2014 i volumi di traffico si sono attestati tra i 30 e i 35 milioni di visitatori unici al mese, generando circa 200-250 milioni di pagine consultate al mese e circa 1 miliardo di visite (non uniche) all'anno.

La mission aziendale di diventare leader nel campo dell'Online Talent Acquisition pone Jobrapido davanti a grosse sfide sia lato B2B che B2C; l'offerta commerciale di fornire On Demand Qualified Candidates permette ai clienti, anche tramite un periodo di prova, di usufruire dei migliori strumenti per trovare i candidati desiderati, su richiesta, in qualsiasi momento. Per il jobseeker vuol dire che Jobrapido può essere un alleato con cui affrontare le sfide future per avere il miglior lavoro possibile. L'esigenza di analizzare i dati e di creare un prodotto che si basi su questi è sempre uno dei punti di forza di Jobrapido; le soluzioni Big Data permettono di rendere sempre più sofisticati i modelli analitici, con l'obiettivo di segmentare meglio la base utenti e le posizioni realmente aperte, al fine di erogare un livello di servizio di molto superiore sia ai jobseeker che ai clienti o recruiter. Le piattaforme di raccolta dati richiedono di porre particolare attenzione alla policy sulla privacy degli utenti e dei loro dati sensibili, per questo motivo Jobrapido opera in pieno rispetto della legge Italiana in vigore in tema di gestione del dato (cookie policy).

Caso particolare di applicazione dei Big Data è quello della profilazione dei job seeker. A queste persone Jobrapido richiede solamente l'indirizzo mail per accedere al servizio offerto, pertanto gli utenti rimangono anonimi e non riconoscibili. Jobrapido per poter essere capace di offrire loro il lavoro più opportuno deve capire però ulteriori caratteristiche. Per fare questo è stato svolto un lavoro di exploration per identificare delle etichette utilizzabili ai fini della classificazione e quindi della profilazione. A tale scopo, inizialmente, sono stati identificati cluster di utenti omogenei sulla base di sei indici ricondotti opportunamente combinati in due variabili target mediante PCA: la frequenza di utilizzo del servizio da parte dell'utente e l'intensità di utilizzo.

Sulla base dei valori assunti dai sei diversi indici sono stati determinati quattro diversi gruppi:

1. Alta frequenza e alto utilizzo: persone che utilizzano il servizio in maniera intensa e duratura;
2. Media/bassa frequenza e alto utilizzo: persone che hanno un livello di regolarità nella visita al sito inferiore, ma un tasso di utilizzo per accesso paragonabile;
3. Media/bassa frequenza e medio/basso utilizzo: persone che accedono poco al servizio e lo utilizzano mediamente;
4. Basso frequenza e basso utilizzo: persone che usano il sito per tempi brevi e che dopo averlo visitato una volta, tendono a ritornarvi dopo un mese o più.

In seguito, l'algoritmo di classificazione ha considerato altri attributi relativi agli avvenimenti che si verificano quando l'utente accede e naviga sul sito, individuando 7 diversi modelli ad alta accuratezza dai quali sono state estratte le regole di classificazione.

Relativamente all'architettura Jobrapido sta utilizzando un approccio da una parte "tradizionale", con un *Datawarehouse* in cui i dati grezzi presenti nei log file vengono elaborati da processi ETL standard per estrarre informazioni, che vengono poi caricati su una piattaforma di Analytics; dall'altra risulta essere innovativo grazie all'introduzione di una nuova *piattaforma* (sull'impronta definita della *Lambda Architecture*). Questa nuova piattaforma, caratterizzata da rapidità e flessibilità, consente di memorizzare su uno storage interno una mole di dati (strutturati e non) in forte crescita di volume di elaborarli mettendo insieme diverse informazioni per fornire risposte utili con elevata rapidità. Questa soluzione supporta Jobrapido nel comprendere i bisogni degli utenti e nel modulare la sua offerta adattandola a ciascun mercato. Con l'adozione di un approccio misto, si sono evoluti anche i team che seguono i progetti Big Data, diventando più eterogenei e con nuove competenze tecniche: mentre prima il team era composto principalmente da ingegneri del software e da analisti, oggi sono state inserite anche figure quali DevOps e Data Scientist.

Un fattore alla base del successo del progetto è la semplicità del servizio: ci sono infatti solamente due barre di ricerca, "che lavoro cerchi" e "dove lo cerchi", che gli utenti devono completare senza fornire necessariamente informazioni aggiuntive. Pur essendo semplice nella relazione con l'utente, il servizio riesce a soddisfare le richieste dei clienti, garantendo la miglior esperienza sul sito, anche grazie all'analisi dei dati di navigazione.

L'azienda, grazie all'elevato livello di servizio offerto, è riuscita a raggiungere il massimo numero di sottoscritti, più di 60 milioni, con conseguenti vantaggi sul fatturato. Inoltre, grazie

all'adozione di un approccio doppio, tradizionale e innovativo, Jobrapido è in grado di memorizzare un numero sempre maggiore di dati, di fare delle analisi sempre più rapide e accurate e di velocizzare il processo di sviluppo dei prodotti. Quest'ultimo aspetto gli consente di incrementare l'innovazione in termini di speed-to-market e di profilare i clienti più precisamente in modo da poter offrir loro le proposte più adeguate.

Una criticità affrontata dall'azienda ha riguardato la capacità di gestire le grosse quantità di dati a disposizione, dato che al mese 100 milioni di utenti visitano il sito generando un grande traffico di dati. Altre difficoltà incontrate riguardano i costi e i tempi dell'elaborazione di alti volumi di dati da mantenere bassi e la scelta degli algoritmi da applicare per le diverse elaborazioni.

L'azienda si sta ponendo per il futuro una serie di sfide. Il primo obiettivo è rappresentato dall'arricchimento del dato sul cliente, ovvero dall'estrazione di valore da quello che è già disponibile sul web. Il fine è migliorare l'esperienza dell'utente, e di conseguenza i risultati raggiungibili, non solo attraverso la tassonomia degli annunci di Jobrapido ma anche in base ai comportamenti di navigazione sul Web e alle informazioni sull'utente diffuse in rete. L'azienda intende anche cominciare ad utilizzare i *second party Data*, da affiancare ai *first party Data*, per migliorare la selezione dei profili di coloro che cercano lavoro forniti alle aziende. Jobrapido sta infine valutando la possibilità di modificare il servizio offerto aggiungendo la funzionalità delle pagine personali per gli utenti, in cui indicare le informazioni private ritenute utili dall'utente stesso per trovare lavoro.

## 6.7 LASTMINUTE.COM

Lastminute.com è un gruppo europeo leader nel settore del turismo online che si pone l'obiettivo di semplificare il processo di selezione e acquisto di viaggi e attività di svago attraverso la tecnologia. Si tratta di un gruppo, nato nel 2004 in Italia con il primo sito di comparazione di voli lowcost Volagratis.com, che poi è diventato BravoflyRumbo Group e oggi è lastminute.com Group con sede a Chiasso, in Svizzera.. In particolare i brand facenti parte del Gruppo sono lastminute.com, bravofly.com, volagratis.com, rumbo.com e il metamotore Jetcost. L'azienda è presente in 35 Paesi attraverso siti web e applicazioni mobile disponibili in quindici lingue, raccogliendo oltre 10 milioni di clienti. Al suo interno lavorano più di 1000 dipendenti. E' quotato alla borsa di Zurigo e nel 2014 ha raggiunto i 147 milioni di euro di ricavi.

L'idea per l'implementazione di un progetto Big Data è nata dalla volontà di far leva sull'enorme base di dati del gruppo, considerando che il sito web conta più di quindici milioni di visualizzazioni giornaliere, per poter ottenere un vantaggio competitivo rispetto ai competitor.

I dati a disposizione del Gruppo vengono generati da diversi canali, inclusi i *metasearch*, che permettono di impostare le ricerche a partire dalle richieste dei potenziali clienti nei diversi motori di ricerca e dai software Over the Air. Con i primi il Gruppo ha a disposizione informazioni sulle ricerche e quindi sulle preferenze e sui comportamenti di navigazione degli utenti, mentre con i secondi è possibile conoscere i comportamenti d'acquisto e le informazioni relative ai pagamenti. I dati risultano sia di natura strutturata che destrutturata, e provenienti sia da fonti interne che esterne.

Gli strumenti analitici utilizzati permettono di conoscere il cliente e di comprenderne i suoi bisogni molto rapidamente, estraendo *insight* che possano poi essere trasformati in azioni rapide ed efficaci, con offerte personalizzate e garantendo le migliori condizioni al consumatore.

I progetti di Big Data, sviluppati in collaborazione con Quantyca, sono realizzati e portati avanti da un team composto da *Data scientist* presenti nell'area di BI & Big Data Analytics, distribuito geograficamente tra le diverse sedi sul territorio europeo.

La criticità principale che il Gruppo ha dovuto gestire è il passaggio da un mondo "volo-centric" di BravoflyRumbo Group ad uno più orientato ai pacchetti come quello di lastminute.com, che presentano un'offerta e un approccio molto differente. Il primo è infatti caratterizzato da un mercato ad alto volume e basso margine che richiede quindi un'ottimizzazione molto spinta in cui l'analisi automatica del dato è molto importante; nel secondo i volumi sono ridotti ma con margini superiori, trattandosi principalmente di crociere e pacchetti vacanza, e si richiede una forte personalizzazione e uno studio attento della dinamica di ricerca e soddisfazione dei clienti. La difficoltà è stata perciò nell'interpretazione di due mentalità ed esigenze diverse e nella fusione di queste per sviluppare un progetto che possa estrarre fattori di successo da entrambi i mondi e dalla combinazione dei due. Un'ulteriore criticità affrontata è quella relativa alla mancanza di coesione iniziale tra il team di BI & Big Data Analytics e quello degli sviluppatori, per via delle competenze eterogenee sia tecnologiche che di business.

In futuro lastminute.com Group vuole sviluppare un algoritmo *multi-armed bandit*, che verrà implementato nel 2016 e che prevede la configurazione di diversi profili contemporaneamente

attivi sul sito e la deviazione del traffico *real-time*, per poi bilanciare gli alberi di decisione in modo che, a seconda del momento, venga deciso quale delle configurazioni possa essere vincente in termini di margini e di conversione. Un secondo progetto riguarderà i *log Analytics* e le attività basate sulle geo-localizzazione: lastminute.com possiede già un motore GSL (*Global Location Services*) in grado di individuare cluster geografici specifici del mercato turistico basandosi sull'area di servizio degli hotel.

## 6.8 VENERE.COM

Venere.com è un'azienda italiana leader nel settore delle prenotazioni alberghiere online che offre 200.000 sistemazioni, dagli hotel ai B&B, appartamenti e agriturismi, in più di 30.000 località sparse per tutto il mondo. Venere agisce da intermediario tra coloro che domandano un'accomodation e coloro che la offrono. Nel settembre 2008 viene acquisita da Expedia Inc, compagnia Statunitense specializzata nell'online travel.

L'azienda, specialista del Sud Europa, ha un modello di business tradizionale: l'utente prenota una sistemazione sul sito, al momento del check-out paga l'importo totale e a questo punto Venere.com riceve una commissione, pari ad una percentuale del valore di prenotazione pagato, dall'offerente dell'accomodation.

La considerazione di utilizzare soluzioni di Big Data nasce dalla volontà di migliorare le esperienze di navigazione e d'acquisto dei consumatori per massimizzare i volumi di prenotazione. Risulta un tema importante in particolare nella parte di CustomerAcquisition, considerata la significativa mole di dati raccolti per fare analisi e per fare Marketing, provenienti sia da fonti interne quale il sito dell'azienda, che da fonti esterne, come i dati forniti dai partner.

Tra i progetti in ambito Big Data che l'azienda sta portando avanti, ve ne sono due molto significativi.

Il primo riguarda il *Search Engine Optimization (SEO)*, ovvero la relazione tra il sito Venere.com e i motori di ricerca: l'obiettivo è unificare e gestire i dati generati da queste interazioni, in modo da stabilire il rapporto causa-effetto tra quest'ultime e le performance di business. Le ricerche e le interazioni da parte degli utenti sui motori di ricerca generano un'enorme quantità di dati: il progetto prevede di raccogliere tutti questi e di organizzare il sito con un'architettura più gerarchica e strutturata in modo da associare i dati alla singola pagina web. In questo modo Venere.com può avere una visione delle performance sia a livello della singola richiesta sia per i

diversi livelli di aggregazione e può capire se ci sono correlazioni significative tra di essi. In questo modo può stabilire quali siano le azioni più opportune da implementare per migliorare e ottimizzare le prestazioni.

Il secondo progetto riguarda il *searchengine Marketing (SEM)*, ovvero gli annunci a pagamento sui motori di ricerca. L'approccio seguito è di *destination-management*, volto all'identificazione dei *point of sale* e dei *point of supply* per capire dove sono le opportunità di business e come queste evolvono nel tempo. Si parte dai dati relativi alla quantità di ricerche effettuate riguardo Venere.com e da quelli relativi a come queste ricerche evolvono nel tempo, tenendo in considerazione le performance di business storiche e stagionali. L'obiettivo è mettere in relazione questi due tipi di informazione per agire sulle campagne facendo pulizia del traffico e interagendo con il *bidding*.

Queste soluzioni, basate sull'utilizzo di algoritmi predittivi e simulazioni, sono caratterizzate da un'elevata efficacia, che comporta relazioni più proficue con search e metasearch e una maggiore visibilità di Venere.com su tali motori di ricerca. La raccolta e l'analisi di tutti i dati, permette inoltre di conoscere meglio il cliente e di proporgli l'offerta più opportuna, al miglior prezzo e al momento giusto. La segmentazione per interessi, click, prenotazioni effettuate e navigazione nel sito comporta comunicazioni più targetizzate e allo stesso tempo flessibili. Questo si traduce in un miglioramento delle prestazioni quali il tasso di conversione degli utenti e i ricavi dalle prenotazioni, essendo le persone maggiormente attratte dal sito web.

Una delle maggiori criticità che l'azienda ha dovuto affrontare è quella relativa alla complessità e all'affidabilità della gestione di una moltitudine di dati sparsi tra le diverse realtà dei partner e che non consentono una stima affidabile dei ritorni.

Una sfida futura che l'azienda è in procinto di affrontare è la customizzazione dei *tools* e degli algoritmi a seconda del partner con cui avviene l'interazione, poiché ognuno di questi si relaziona con il sito dell'agenzia e con gli utenti in modalità differenti. Infine Venere.com desidera analizzare i diversi canali di comunicazione e vendita per poter approfondire l'interazione esistente tra questi.





## BIBLIOGRAFIA

- Alpine Data Labs. (2015). *Alpine Data Labs: Solutions*. Tratto il giorno Novembre 16, 2015 da Sito Web. Alpine Data Labs: <http://alpinedata.com/products/>
- Alton L. (2015, Luglio 01). *Blog: 10 Ways Location-Based Marketing Will Evolve in 2015*. Tratto il giorno Ottobre 01, 2015 da Sito Web: Huffpost business: [http://www.huffingtonpost.com/larry-alton/10-ways-locationbased-mar\\_b\\_6423832.html](http://www.huffingtonpost.com/larry-alton/10-ways-locationbased-mar_b_6423832.html)
- Ariker M. (2014, Febbraio 04). CMOs e CIOs need to get along to make Big Data work. *Harvard Business Review* .
- Ariker M., M. T. (2013). Five Roles You Need on Your Big Data Team. *Harvard Business Review* .
- Baker W., K. D. (2014, Giugno). Using Big Data to make better pricing decisions. *McKinsey Quarterly* .
- Baumgartner T., H. H. (2011). Find Big Growth in Big Data. In H. H. Baumgartner T., *Sales Growth. Five proven strategies from world's sales leader* (p. 41).
- Bellini L. (2012). Come misurare i Social media, un modello di riferimento. *digitalmMarketinglab* .
- Bernice P. (2013). The emergence of "big data" technology and analytics. *Journal of Technology Research* , 1-6.
- Bhandari R., S. M. (2014, Giugno). Using marketing analytics to drive superior growth. *McKinsey Quarterly* .
- Bikar P. (2015, giugno 01). *Big Data in the Public Sector: Turning Bytes into Information Valuable to Society* .Tratto il giorno settembre 19, 2015 da Sito Web: Cisco Blogs: <http://blogs.cisco.com/government/public-sector-big-data>
- BIT BANG. (2015, Luglio 15). *Analytics platform: il futuro della social media analytics*. Tratto il giorno Settembre 30, 2015 da Sito Web: BIT BANG: <http://www.social-media-analytics.it/social-media-analysis/il-futuro-della-social-media-analytics/>
- BIT BANG. (2014, Maggio 29). *Analytics Platform: Social Media Intelligence Framework: una guida nel panorama della Social Media Analysis*. Tratto il giorno Settembre 30, 2015 da Sito Web: BIT BANG: <http://www.social-media-analytics.it/social-media-analysis/social-media-intelligence-framework-una-guida-nel-panorama-della-social-media-analysis/>
- BIT BANG. (2015, Gennaio 23). *Facebook: social analytics per la competitive intelligence*. Tratto il giorno Settembre 30, 2015 da Sito Web: BIT BANG: <http://www.social-media-analytics.it/social-media-analysis/social-analytics-per-la-competitive-intelligence/>
- BIT BANG. (2014, Novembre 07). *Linee guida: http://www.social-media-analytics.it/social-media-analysis/lidentificazione-dei-social-media-influencer/*. Tratto il giorno Settembre 30,

2015 da Sito Web: BIT BANG: <http://www.social-media-analytics.it/social-media-analysis/identificazione-dei-social-media-influencer/>

BIT BANG. (2015, Gennaio 19). *Linee guida: il business value della social intelligence*. Tratto il giorno Settembre 30, 2015 da Sito Web: BIT BANG: <http://www.social-media-analytics.it/social-media-analysis/social-intelligence-business-value/>

BIT BANG. (2014, Dicembre 23). *Social Media Analysis: Come si evolverà la social intelligence nel 2015?* Tratto il giorno Settembre 30, 2015 da Sito Web: BIT BANG: <http://www.social-media-analytics.it/social-media-analysis/come-si-evolvera-la-social-intelligence-nel-2015/>

BIT BANG. (2015, Luglio 24). *Social Media Monitoring. Visual listening: il nuovo campo della social intelligence*. Tratto il giorno Settembre 30, 2015 da Sito Web: BIT BANG: <http://www.social-media-analytics.it/social-media-monitoring/visual-listening-il-nuovo-campo-della-social-intelligence/>

Buckley J. (2015, Gennaio 14). *Blog: Big Data and Customer Micro-Segmentation: Applications in Media and E-Commerce*. Tratto il giorno Ottobre 14, 2015 da Sito Web: Qubole: <https://www.qubole.com/blog/big-data/big-data-customer-microsegmentation/>

Casali A. (2015, Luglio 27). *Digital Marketing. Mobile, video, big data e social: internet trasforma la pubblicità*. Tratto il giorno ottobre 2, 2015 da Sito Web Digital4: [http://www.digital4.biz/digital4/mobile-video-big-data-e-social-internet-trasforma-la-pubblicita\\_43672155427.htm](http://www.digital4.biz/digital4/mobile-video-big-data-e-social-internet-trasforma-la-pubblicita_43672155427.htm)

Casali A. (2015, Luglio 27). *Digital Marketing. Mobile, video, big data e social: internet trasforma la pubblicità*. Tratto il giorno ottobre 2, 2015 da Sito Web 4Digital : [http://www.digital4.biz/digital4/mobile-video-big-data-e-social-internet-trasforma-la-pubblicita\\_43672155427.htm](http://www.digital4.biz/digital4/mobile-video-big-data-e-social-internet-trasforma-la-pubblicita_43672155427.htm)

Chen H., C. R. (2012, Dicembre). Business Intelligence and Analytics: from Big Data to Big Impact. *MIS QUARTERLY* , p. 1165-1188.

Court D. (2015). Getting big impact from big data. *McKinsey Quarterly* .

Dell. (2015). *Solution & Services: Big Data*. Tratto il giorno Novembre 16, 2015 da Sito Web Dell: <http://www.dell.com/en-us/work/learn/business-intelligence-big-data?~ck=mn>

Edelman D., S. M. (2015, Ottobre). The new consumer decision journey. *McKinsey Quarterly* .

Fanderl H., M. N. (2014, Gennaio 23). Do you have the right stuff? Three steps for turning Big Data chaos in customer experience gold. *McKinsey Quarterly* .

Fiorenza P., P. A. (2014). *The Next Frontier of Public Sector Innovation: Big Data*. GovLoop.

Gartner. (2015). *Magic Quadrant for Advanced Analytics Platforms*.

Gartner. (2015). *Magic Quadrant for BI Intelligence and Analytics Platform*

Generali. (2015, Luglio 23). *Media. Comunicati Stampa 2015: Generali acquisisce MyDrive Solutions, società leader nelle tecnologie per l'assicurazione auto*. Tratto il giorno Novembre 24, 2015 da Sito Web Generali: <http://www.generali.com/it/media/press-releases/all/2015/Generali-acquires-MyDrive-Solutions--a-leading-company-in-motor-insurance-technologies.html>

Germann F., L. G. (2012). Performance implications of deploying marketing analytics. *Intern. J. of Research in Marketing* , 114-128.

Hazan E., B. F. (2013, Agosto). Leveraging big data to optimize digital marketing. *McKinsey Quarterly* .

Holtz D. (2014, novembre 07). *8 Skills You Need to Be a Data Scientist*. Tratto il giorno settembre 22, 2015 da Sito Web UDACITY: <http://blog.udacity.com/2014/11/data-science-job-skills.html>

Hu Y., Y. T. (2012). Discovering valuable frequent patterns based on RFM analysis without. *Knowledge-Based Systems* , 76–88.

IBM. (s.d.). *Big Data and Analytics. Marketing: Acquisire, far crescere e mantenere i clienti*. Tratto il giorno Settembre 26, 2015 da Sito Web: IBM: <http://www.ibm.com/big-data/it/it/big-data-and-analytics/marketing.html>

IBM. (s.d.). *What is a data scientist?*Tratto il giorno settembre 22, 2015 da Sito Web: IBM: <http://www-01.ibm.com/software/data/infosphere/data-scientist/>

Jeske M., G. M. (2013). *BIG DATA IN LOGISTIC. A DHL perspective on how to move beyond the hype*. DHL Customer Solutions & Innovation.

Kayyali B., K. D. (2013, Aprile). The big-data revolution in US health care: Accelerating value and innovation. *McKinsey Quarterly* .

Kerschberg B. (2014, dicembre 01). *How Big Data and Business Intelligence are changing the Manufacturing Industry*. Tratto il giorno settembre 15, 2015 da Sito Web: APPIRIO: <http://appirio.com/category/business-blog/2014/12/big-data-business-intelligence-changing-manufacturing-industry/>

Lau R., L. C. (2014). Social analytics: Learning fuzzy product ontologies for aspect-oriented. *Decision Support Systems* , 80-94.

Lee J., K. H. (2014). Service innovation and smart analytics for Industry 4.0 and big data environment. *ScienceDirect* .

Lühr P., M. R. (2013, Febbraio). Name your price: The power of Big Data and analytics. *McKinsey Quarterly* .

Manyika J., C. M. (2011). *Big Data: The Next Frontier for Innovation, Competition, and Productivity*.

MarketingLand. (2014, Febbraio 10). *Mobile Marketing. Infographic: How Location-Based Marketing Delivers Ads To In-Store Shoppers*. Tratto il giorno Ottobre 01, 2015 da Sito Web: MarketingLand: <http://marketingland.com/location-based-marketing-shoppers-73574>

McAfee A., B. E. (2012). Big Data: The Management Revolution . *Harward Business Review* , 60-68.

Michael K., M. K. (2013). Big Data: New Opportunities and New Challenges. *Browse Journal & Magazines* , 22-24.

Miller K. (2011-2012, inverno). Big Data Analytics in Biomedical Research. *Biomedical Computation Review* , p. 14-21.

O'Brien B. (2014, Aprile 30). *Blog. Mobility. Explosive Growth in Mobility and Location-based Marketing*. Tratto il giorno Ottobre 01, 2015 da Sito Web: Cisco: <http://blogs.cisco.com/wireless/explosive-growth-in-mobility-and-location-based-marketing>

Offsey S. (2015). *Blog: Micro-Segmentation in the Age of Big Data*. Tratto il giorno Ottobre 14, 2015 da Sito Web Marketbuildr: <http://marketbuildr.com/blog/segmentation-in-the-age-of-big-data/>

Opresnik D., T. M. (2015). The value of Big Data in servitization. *International Journal of Production Economics* .

OptiMine Software. (2015). *OptiMine: Products*. Tratto il giorno Novembre 16, 2015 da Sito Web OptiMine: <http://optimine.com/>

Osservatorio Big Data Analytics & Business Intelligence. (maggio 2015). Introduzione: I Big Data come leva strategica nel Retail. *Workshop: I Big Data come leva strategica nel Retail*. Milano.

Osservatorio Big Data Analytics & Business Intelligence. (2014). Risultati Ricerca 2014 Osservatorio Digital Innovation. *Big Data: un mercato in cerca d'autore*. Milano.

Osservatorio Big Data Analytics & Business Intelligence. (2015). Risultati Ricerca 2015 Osservatorio Digital Innovation. *Big Data: da Data Insight a Data Driven Strategy*

Prexion. (2015). *Predixion: Product*. Tratto il giorno Novembre 16, 2015 da Sito Web Predixion: [http://www.predixionsoftware.com/overview\\_45](http://www.predixionsoftware.com/overview_45)

Raghupathi W., R. V. (2014). Big data analytics in healthcare: promise and potential. *Health Information Science and Systems* .

RetailNext. (2015). *How It Works: RetailNext technology works through a series of steps*. Tratto il giorno Ottobre 02, 2015 da Sito Web: RetailNext: <http://retailnext.net/how-it-works/>

Sage. (2013). *Better Inventory Management. Big Challenges, Big Data, Emerging Solutions*. Sage Software.

SAS. (2015). *Prodotti e Soluzioni: Advanced Analytics Software*. Tratto il giorno Novembre 16, 2015 da Sito Web SAS: [https://www.sas.com/it\\_it/software/customer-intelligence.html](https://www.sas.com/it_it/software/customer-intelligence.html)

- Seret A., V. T. (2012). A new SOM-based method for profile generation: Theory and an application. *European Journal of Operational Research* , 199–209.
- ShopperTrack. (2015). *Solutions: Retail Profitability, Improved. Take your business to the next level*. Tratto il giorno Ottobre 02, 2015 da Sito Web: ShopperTrack: <http://www.shoppertrak.com/retail-analytics/>
- Sun J., R. C. (2013). Big Data Analytics for Healthcare. *SIAM International Conference on Data Mining*. Austin, TX.
- Tafner C. (2013, Novembre 28). *Marketing & ADV blog. Real Time Bidding: Cos'è? Come funziona? Chi vince?* Tratto il giorno Ottobre 2, 2015 da Sito Web: Marketing Caffè : <http://www.marketingcaffe.com/it/real-time-bidding-rtb/>
- Tarka P., Ł. M. (2014). Decision Making in Reference to Model of Marketing Predictive Analytics. *Management and Business Administration. central europe* , p. 60-69.
- Teerlink M., S. P. (2014). *The new hero of big data and analytics. The Chief Data Officer*. IBM Global Business Services.
- Teradata. (2015, Febbraio). *Forbes Insights: Teradata 2015 Global Data-Driven Marketing Survey*. Tratto il giorno Settembre 27, 2015 da Sito Web: Forbes: [http://images.forbes.com/forbesinsights/StudyPDFs/Teradata-Data\\_Driven\\_Marketing-REPORT.pdf](http://images.forbes.com/forbesinsights/StudyPDFs/Teradata-Data_Driven_Marketing-REPORT.pdf)
- Van Rijmenam M. (2015, Maggio 08). *Chief Data Officer Profile*. Tratto il giorno Settembre 21, 2015 da Sito Web DATAFLOQ: <https://datafloq.com/read/job-description-chief-data-officer/196>
- Van Rijmenam M. (2015, maggio 06). *Data Scientist Profile*. Tratto il giorno settembre 21, 2015 da Sito Web DATAFLOQ: <https://datafloq.com/read/job-description-big-data-scientist/197>
- Vercellis C. (2005). *Business Intelligence*. Milano: McGraw Hill.
- Visaul IQ. (2015). *Visual IQ: Products*. Tratto il giorno Novembre 16, 2015 da Sito Web Visual IQ: <http://www.visualiq.com/products/iq-deploy>
- Wikipedia. (2015, Ottobre 28). *Magic Quadrant*. Tratto il giorno Novembre 16, 2015 da Sito Web Wikipedia: [https://en.wikipedia.org/wiki/Magic\\_Quadrant](https://en.wikipedia.org/wiki/Magic_Quadrant)
- Zanotti L. (2015, Settembre 10). *Product Lifecycle Management. RFID: dall'identification of things alla Internet of Things, ecco cosa sta cambiando*. Tratto il giorno 1 Ottobre, 2015 da Sito Web: Digital4: [http://www.digital4.biz/iot/news/rfid-dall-identification-of-things-alla-internet-of-things-ecco-cosa-sta-cambiando\\_43672155903.htm](http://www.digital4.biz/iot/news/rfid-dall-identification-of-things-alla-internet-of-things-ecco-cosa-sta-cambiando_43672155903.htm)
- Zinck B. M. (2015, Agosto 18). *Gartner's 2015 Hype Cycle for Digital Marketing*. Tratto il giorno Novembre 18, 2015 da Sito Web Digital Tech Diary: <http://digitaltechdiary.com/gartners-2015-hype-cycle-for-digital-marketing/2241/>

