



Dipartimento di Impresa e Management  
Cattedra di Economia e Gestione delle Imprese

## ANALISI DEL FENOMENO DEI BIG DATA

RELATORE

Prof. Luca Pirolo

CANDIDATO

Davide Signore

Matricola 182441

ANNO ACCADEMICO

2015/2016



INTRODUZIONE.....	4
1. IL MONDO DEI BIG DATA .....	6
1.1 LA QUANTITA' DELLE INFORMAZIONI .....	6
1.2 COSA SONO I BIG DATA? .....	9
1.3 LA CATENA DEL VALORE.....	13
1.3.1 IL FUTURO DEI BIG DATA.....	16
1.3.2 GLI INTERMEDIARI DEI DATI .....	18
1.4 L'ALTRA FACCIA DELLA MEDAGLIA .....	21
2. APPLICAZIONE DEL FENOMENO DEI BIG DATA .....	25
2.1 IL FUTURO MERCATO DEI BIG DATA .....	25
2.1.1 ANALISI SWOT .....	28
2.2 DATA-DRIVEN BUSINESS MODEL.....	30
2.2.1 QUADRO TEORICO PER LA CREAZIONE DI UN DBBM .....	34
2.3 NUOVI RUOLI NECESSARI GESTIRE UN BIG DATA TEAM.....	37
2.4 EFFETTI SUL SETTORE DELLA VENDITA AL DETTAGLIO.....	39
CONCLUSIONE.....	53
BIBLIOGRAFIA/SITOGRAFIA.....	56

# INTRODUZIONE

Il presente lavoro esplora il mondo dei big data, fenomeno che sta modificando il focus all'interno delle aziende. Essendo quello dei big data un argomento che non ha trovato ancora una definizione specifica, la ricerca ha una natura di carattere esplorativo e cerca di analizzare e studiare le caratteristiche del fenomeno da più punti di vista, cercando di proporre sempre spunti pratici interessanti. L'obiettivo di questa tesi è, offrire un valido supporto per la comprensione del fenomeno dei big data e la possibilità di ottenere vantaggi competitivi. L'evoluzione tecnologica risulta essere il presupposto necessario al fine di rendere concreto il fenomeno. Questa infatti è lo strumento di supporto per la raccolta, analisi e la gestione di un numero via via crescente di dati. Solo attraverso la tecnologia sarà possibile favorire l'efficienza operativa, le performance produttive, le relazioni con la clientela, la trasparenza, l'innovazione nei prodotti/servizi e nei modelli di business.

Vediamo passo per passo come si sviluppa l'analisi.

La prima parte dell'elaborato introduce i Big Data e ne descrive il contesto, il cambiamento della nostra società che ha portato alla nascita del fenomeno. Analizza le caratteristiche, le "4V" (volume, varietà, velocità e valore), mettendo in risalto i vantaggi che l'utilizzo dei dati può apportare alle aziende. Ma esistono anche i lati oscuri dei big data infatti pur offrendo molte opportunità, non sono esenti da rischi. Il primo capitolo si conclude infatti con delle riflessioni legate alle problematiche della privacy. Il secondo capitolo si apre con un'analisi dell'utilizzo dei dati nel contesto aziendale, esplora quelle che sono le applicazioni strategiche da parte delle aziende stesse attraverso grafici ed informazioni. Segue un'analisi Swot, che ha lo scopo di esporre i punti di forza e le opportunità, ma anche gli svantaggi e le

minacce che seguono dall'utilizzo dei big data. L'applicazione di tale fenomeno all'interno delle imprese comporta diverse conseguenze; in primo luogo infatti i business model delle imprese dovranno adeguarsi, in modo tale da poter trarre il massimo vantaggio dalla raccolta e quindi dall'utilizzo dei dati stessi. Nasce quindi la necessità di costruire nuovi business model, in particolare il "data driven business model", che vede nelle informazioni, una delle risorse fondamentali per produrre vantaggio competitivo. Un'altra conseguenza sarà la necessità di valorizzare il fattore umano, necessario al fine di utilizzare al meglio i dati e farli diventare una risorsa all'interno dei business. Si parlerà quindi dei nuovi ruoli necessari per poter gestire una sempre più numerosa mole di dati. Al fine di mostrare nella pratica i vantaggi che derivano dall'utilizzo dei big data, segue un'analisi specifica del settore della vendita al dettaglio, ricerca svolta dal McKinsey Global Institute, che identifica quelli che potranno essere gli effetti sul settore e le possibili barriere allo sviluppo. Il capitolo si chiude con l'analisi di una delle aziende che sta sfruttando nel migliore dei modi i big data, Netflix. Essa infatti genera una vastissima mole di dati e riesce a tradurli in informazioni nel modo migliore possibile, riuscendo in questo modo ad aggiudicarsi una fetta di mercato sempre maggiore, come dimostrato dalle informazioni esposte.

# 1. IL MONDO DEI BIG DATA

## 1.1 LA QUANTITA' DELLE INFORMAZIONI

Oggi il mondo è, come mai prima d'ora, sommerso da informazioni e queste, soprattutto negli ultimi vent'anni sono cresciute nel numero sempre più rapidamente. Cercare di quantificare dati ed informazioni che ci circondano, risulta molto complicato.

Un risultato soddisfacente è stato raggiunto da Martin Hilbert, professore presso la "School for Communication and Journalism", Southern California.

Hilbert ha tentato di quantificare tutte le informazioni e dati prodotti nel mondo dal 1986 al 2007, includendo nella sua ricerca dati provenienti da fonti analogiche e digitali: DVD's, e-mail, foto, musica, come anche videogames, telefonate e trasmissioni radio foniche.

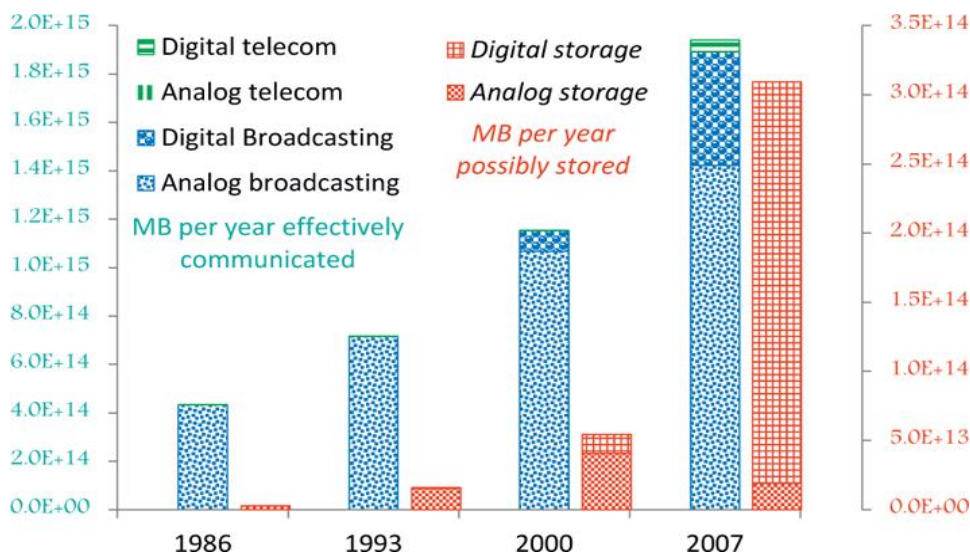


FIGURA 1 HOW MUCH INFORMATION IS THERE IN THE "INFORMATION SOCIETY?"

Come riportato dal grafico (fig.1), nel 2007 sono stati archiviati circa 300 exabyte di dati. Si può notare l'incredibile rapidità con cui i dati sono cresciuti nel tempo e la loro digitalizzazione: mentre nel 2000 solo il 25% dei dati immagazzinati nel mondo era in forma digitale, nel 2007 invece la situazione si è capovolta, risultando così solo un 7% dei dati in forma analogica e la restante parte in forma digitale.

Dall'analisi svolta Hilbert conclude che, i dati si espandono raddoppiando ogni 3 anni a riprova di questo, per il 2013 la quantità di informazioni immagazzinata è stata stimata in 1200 exabyte<sup>1</sup>.

L'aumento esponenziale delle informazioni può essere dimostrato dai seguenti dati: Google processa più di 24 petabyte di dati al giorno; su Facebook si caricano ogni giorno più di 10 milioni di nuove fotografie e gli iscritti cliccano "Mi piace" quasi tre miliardi di volte al giorno in questo modo l'azienda può analizzare e capire le preferenze degli utenti<sup>2</sup>. Intanto su Twitter il numero di messaggi cresce circa del 200 per cento all'anno<sup>3</sup>.

Come si può notare quindi il fenomeno è in costante crescita, secondo Peter Norvig, esperto di intelligenza artificiale di Google, il cambiamento quantitativo produrrà un cambiamento qualitativo. Dal momento che un film è radicalmente diverso da una fotografia e ancor di più da un dipinto; la stessa cosa si potrà dire dei big data: modificandone la quantità, ne modifichiamo l'essenza<sup>4</sup>.

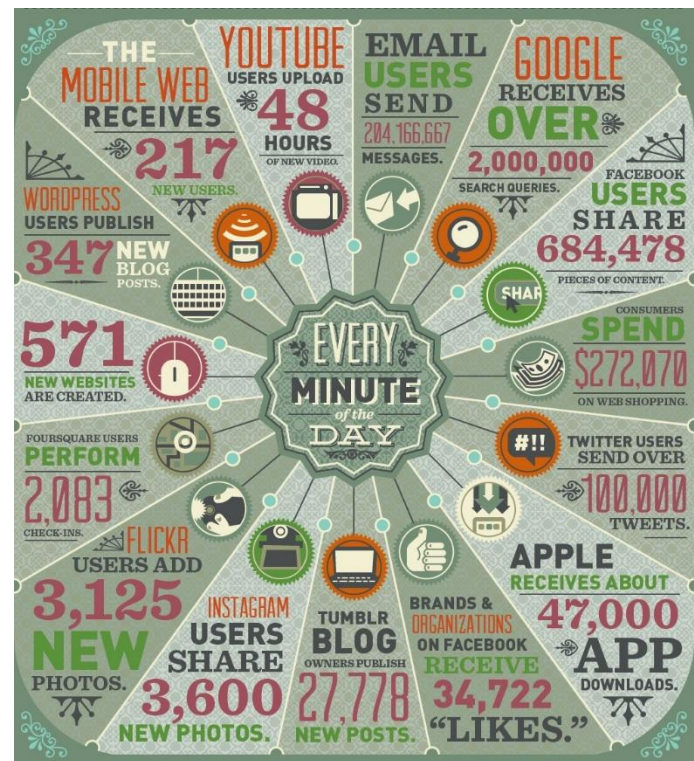


FIGURA 2 : "QUANTI DATI POSSONO ESSERE GENERATI OGNI MINUTO?"

Si può affermare che vi sono tre condizioni fondamentali alla base dei big data interconnesse tra loro: l'aumento della disponibilità delle informazioni, il miglioramento della capacità di elaborazione e di conservazione dei dati stessi,

<sup>1</sup> Hilbert, M., How much information is there in the "information society"? Significance, 9(4), 8–12.

<sup>2</sup> Prospetto IPO di Facebook, <<Form S-1 Registration Statement>>, U.S. Securities and Exchange Commission, 1° febbraio 2012.

<sup>3</sup> T. Geron, Twitter's Dick Costolo: *Twitter Mobile Ad Revenue Beats Desktop on Some Days*, <<Forbes>>, 6 giugno 2012.

<sup>4</sup> Mayer Schönberger V., Cuckier K, *Big Data: una rivoluzione che trasformerà il nostro modo di vivere e già minaccia la nostra libertà*. Milano: Garzanti, 2013

infine la convenienza economica, rispetto al passato, nell'ottenere i due aspetti indicati in precedenza.

Per quanto riguarda l'aumento della disponibilità delle informazioni essa è conseguenza di due fenomeni: la datizzazione e l'internet of things. La prima è considerata il carburante dei big data poiché consente di estrarre e porre in correlazione i dati. Differisce dalla digitalizzazione, la quale consiste nel processo di conversione di elementi analogici in formato digitale. La datizzazione invece si riferisce al processo di conversione in forma quantitativa di un determinato fenomeno, in modo da poterlo tabulare ed analizzare. Tale fenomeno comporterà un cambiamento epocale: la possibilità di utilizzare tutti i dati a disposizione, invece di una serie di campioni ristretti. La datizzazione sta coprendo i più disparati ambiti, nel caso dei social network ad esempio, Facebook trasforma le relazioni in dati, Twitter ha reso possibile la datizzazione dei sentimenti attraverso un sistema piuttosto immediato e LinkedIn le esperienze professionali trasformando le informazioni acquisite in previsioni sul nostro presente e futuro, consigliandoci chi potremmo conoscere o il lavoro che potremmo considerare. L'Internet of things consente invece di raccogliere i dati sullo stato e sul funzionamento degli oggetti, attraverso particolari sensori posti sugli oggetti stessi, un esempio sono i sistemi gps.

La seconda delle tre condizioni è connessa alla legge di Moore relativa alla crescita esponenziale della potenza di calcolo e allo sviluppo dei sistemi di intelligenza artificiale.

Infine si può notare come si sia ridotta notevolmente la spesa economica relativa alla raccolta, all'archiviazione e all'elaborazione dei dati grazie allo sviluppo dei sistemi di intelligenza artificiale e a causa del costo marginale di produzione praticamente nullo provocato dalla trasmissione dei dati digitali.

Quando le posizioni si sono trasformate in dati, aziende che sviluppano applicazioni che raccolgono informazioni sul posizionamento quali per esempio Foursquare, la quale attraverso il "check in" degli utenti nei loro luoghi preferiti, genera profitto sia attraverso programmi di fidelizzazione, dalle raccomandazioni dei ristoranti e da altri servizi legati al posizionamento.



Nel passaggio quindi dagli small ai big data, quando N=tutti non ci si riferisce più semplicemente ad un campione limitato, quindi ovviamente la confusione dei dati con cui confrontarsi aumenterà. Il loro volume giustifica ampiamente la rinuncia all'esattezza che era garantita dai metodi di analisi precedenti, poiché essi non catturano semplicemente una piccola parte del fenomeno, ma il fenomeno nella sua interezza.

## 1.2 COSA SONO I BIG DATA?

*“I dati stanno diventando una nuova materia prima del business: input economico pressoché equivalente al capitale e alla forza lavoro.”* <<Economist>>, 2010

Il termine big data costituisce un concetto astratto, non esiste una definizione esatta ed esclusiva. In primo luogo i big data possono essere definiti in maniera negativa come l'insieme di dati che non può essere rilevato, ottenuto, gestito e analizzato con le tradizionali tecnologie informative e database.

Nonostante quello dei big data sia un fenomeno recente, alcuni ricercatori hanno tentato di esplicitare il fenomeno, esistono quindi diverse definizioni che differiscono per alcuni elementi:

- Nel 2010 Apache Hadoop (framework che supporta applicazioni distribuite con elevato accesso ai dati sotto una licenza libera): *“Set di dati che non può essere catturato, acquisito e gestito da computer generali all'interno di uno specifico ambito.”*

- Nel 2011 La McKinsey and Company (multinazionale di consulenza aziendale) ha definito il fenomeno come la nuova frontiera dell'innovazione, della concorrenza e della produttività. Secondo la compagnia i big data sono infatti quel set di dati che non possono essere acquisiti e gestiti dai classici database. Dalla definizione si ricava che, il volume dei dati non è l'unico criterio da tenere in considerazione, altre due caratteristiche chiavi sono i flussi sempre più crescenti di dati e il management che non potrà più utilizzare i database tecnologici tradizionali.
- Gartner Inc. (società per azioni multinazionale leader mondiale nella consulenza strategica, ricerca e analisi nel campo dell'Information Technology) ha individuato nei big data quelle che sono chiamate le 4V<sup>5</sup>:

MULTIPLI DEL BYTE			
Nome	Simbolo	Multiplo	byte
Kilobyte	kB	10 <sup>3</sup>	1.000
Megabyte	MB	10 <sup>6</sup>	1.000.000
Gigabyte	GB	10 <sup>9</sup>	1.000.000.000
Terabyte	TB	10 <sup>12</sup>	1.000.000.000.000
Petabyte	PB	10 <sup>15</sup>	1.000.000.000.000.000
Exabyte	EB	10 <sup>18</sup>	1.000.000.000.000.000.000
Zettabyte	ZB	10 <sup>21</sup>	1.000.000.000.000.000.000.000
Yottabyte	YB	10 <sup>24</sup>	1.000.000.000.000.000.000.000.000

FIGURA 3: MULTIPLI DEL BYTE

1. **Volume:** con il termine si fa riferimento all'ingente quantità di tutti i tipi di dati generati dalle diverse fonti che non possono essere gestiti dai tradizionali database, ma hanno bisogno di essere organizzate ed analizzate. È facile capire che stiamo trattando informazioni dell'ordine di grandezza dei terabytes per entrare prossimamente nel mondo degli zettabytes (fig.2 mostra l'ordine di grandezza dei byte).

<sup>5</sup> Ibrahim Abaker Targio Hashem, Information Systems journal: The rise of "big data" on cloud computing: Review and open research issues (2014).

Ad esempio il motore di un aereo è capace di generare circa 20 Terabyte di dati ogni ora di volo quindi un volo Milano-New York, con un quadrimotore ne genererà ben 640 TB. È chiaro che la gestione di una simile mole di dati richiede un metodo diverso da quello tradizionale.

2. **Variety**: ovvero i dati di diversa natura che possono essere strutturati e non, raccolti per esempio: tramite smartphones, social networks o ancora tramite transazioni commerciali.
3. **Velocity**: per essere utilizzati questi dati necessitano di un'alta velocità durante la fase del processo di “*data transfer*”, in questo modo i dati potranno essere processati in maniera quasi istantanea garantendo un'elevata funzionalità.

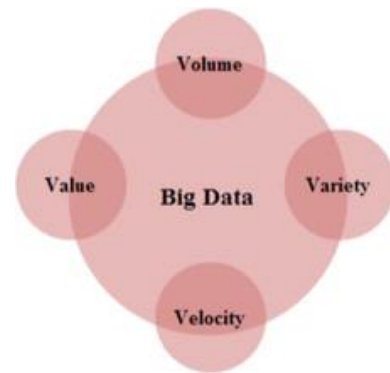


FIGURA 4: 4V'S

Ottenere un vantaggio competitivo infatti significa identificare un'opportunità in minuti o anche a volte qualche secondo prima della concorrenza.

4. **Value**: è considerato l'aspetto più importante dei big data e si riferisce al processo di individuazione di un elevato valore nascosto all'interno di un gran numero di dati, diversi e in rapida crescita. È fondamentale quindi valutare la veridicità e la qualità dei dati affinché possano effettivamente generare nuovo valore.

Se adesso focalizziamo la nostra attenzione sulle imprese che utilizzano i dati e su come si integrano nella catena del valore delle informazioni, si può notare che si stanno affermando tre categorie di imprese, che utilizzano i big data. Esse differiscono tra loro sulla base del valore offerto<sup>6</sup>:

---

<sup>6</sup> Mayer Schönberger V., Cuckier K, *Big Data: una rivoluzione che trasformerà il nostro modo di vivere e già minaccia la nostra libertà*. Milano: Garzanti, 2013

- **Data:** rientrano in questa categoria tutte quelle imprese che hanno la possibilità di accedere alle informazioni in maniera diretta. Fanno parte di questa categoria, Twitter che pur avendo a disposizione nei propri server di una grande quantità di dati, si è rivolta ad aziende indipendenti per darli in licenza ad altre aziende.
- **Competenze:** appartengono a questa categoria tutte quelle società che si focalizzano sulle competenze, appunto e sono principalmente società di consulenza e fornitori di supporti analitici che svolgono nella pratica il lavoro. Rientra in questa categoria la società Teradata, (leader mondiale dell'Enterprise Data Warehousing, che consiste in un archivio informatico di dati di un'organizzazione, per consentire di produrre analisi ai fini decisionali aziendali) che ha svolto e scoprì una correlazione positiva tra previsioni atmosferiche di tempeste ed uragani e consumo di merendine. Infatti si scoprì, analizzando le transazioni passate dei suoi clienti che, in previsione di un uragano, aumentavano gli acquisti di torce elettriche e di merendine da parte dei clienti della grande catena. Tali prodotti, quando le previsioni atmosferiche segnalavano fenomeni atmosferici quali uragani o tempeste, venivano quindi posti all'ingresso dei punti vendita portando quindi ad un incremento delle vendite della multinazionale Walmart.
- **Idee:** in quest'ultima categoria rientrano tutte quelle aziende che sono caratterizzate da una cultura orientata verso i big data, la differenza principale rispetto alle altre categorie è che queste imprese sono contraddistinte dall'utilizzo dei dati in maniera creativa per creare nuove forme di valore, questo rappresenta appunto il loro *core business*. Un esempio pratico è la start up Decide.com, la quale ha iniziato la sua attività nel 2011 come motore di previsione dei prezzi per un a serie di prodotti di largo consumo. I suoi computer recuperavano dati dai vari siti di acquisti online ed esaminava il web alla ricerca di tutte quelle informazioni

che potevano risultare utili in fase d'acquisto e quindi incidere sul prezzo del prodotto. La filosofia di quest'azienda è appunto quella di guidare il consumatore ad acquistare in modo consapevole suggerendo il momento migliore per comprare.



FIGURA 5 DECIDE.COM: ACQUISTO CONSAPEVOLE

Competenze e dati fino ad ora sono quelle caratteristiche che hanno ricevuto maggiore attenzione, oggi però sta prendendo piede una nuova figura professionale: quella del “data scientist”. Esso infatti analizza i dati per fare nuove scoperte. Secondo il chief economist di Google, Hal Varian: “Se vuoi avere successo devi opporre alternative e scarsità a un bene ubiquitario e venduto a basso prezzo [...] i dati sono così largamente disponibili che il bene scarso sono le conoscenze necessarie per estrarne utili indicazioni<sup>7</sup>”. Probabilmente però con l’evoluzione del settore, le risorse, quali i dati, acquisteranno maggior valore a discapito delle competenze mitizzate da Hal Varian.

### 1.3 LA CATENA DEL VALORE

È importante quindi analizzare nel particolare la *chain value* dei big data per capire quella che sarà l’evoluzione delle aziende. Possiamo distinguere tre componenti fondamentali:

---

<sup>7</sup> Mayer Schönberger V., Cuckier K., *Big Data: una rivoluzione che trasformerà il nostro modo di vivere e già minaccia la nostra libertà*. Milano: Garzanti, 2013

- **Detentori dati:** essendo l'essenza dei big data proprio le informazioni, essi rappresentano il punto di partenza dell'analisi della catena del valore. Questi sono coloro i quali hanno l'accesso alle informazioni e le utilizzano direttamente o indirettamente dandole in licenza ad altre aziende. Un esempio è l'azienda ITA Software la quale ha fornito a Farecast (azienda fondata nel 2003 che, come Decide.com, consigliava ai consumatori se e quando acquistare un biglietto aereo. L'azienda è stata acquistata da Microsoft nel 2009 per \$115 milioni<sup>8</sup>) i propri dati per effettuare analisi e previsioni sulle tariffe aeree. ITA Software ha deciso quindi di non sfruttare i dati, come invece ha fatto Farecast acaparrandosi la maggior parte del valore secondario dei dati stessi. Altre imprese nel settore delle carte di credito, ad esempio MasterCard o Visa, si sono posizionate al centro del flusso informativo. Nel caso di MasterCard infatti esiste una divisione specializzata nella raccolta e analisi dei dati "MasterCard Advisors" che appunto analizza 65 miliardi di transazioni in 210 paesi, per trovare tendenze costanti nel comportamento dei consumatori, che poi rivende ad altre aziende<sup>7</sup>. Da queste analisi è risultato ad esempio che se gli automobilisti faranno il pieno tra le quattro e le cinque del pomeriggio, spenderanno tra i 30 ed i 50 dollari all'interno di un supermercato o di un ristorante. Queste informazioni che se fossero analizzate singolarmente non assumerebbero un significato così rilevante, invece in questo modo potranno essere sfruttate da operatori di marketing che potrebbero ad esempio far stampare sul retro degli scontrini della benzina, buoni da utilizzare all'interno di un ristorante o di un supermercato della zona.
- **Data specialists:** sono le aziende che sono in possesso delle competenze e delle tecnologie che servono per effettuare un'analisi complessa, diversamente da come abbiamo visto prima con il caso MasterCard alcune imprese decidono di rivolgersi a cosiddetti specialisti, un esempio è la società

---

<sup>8</sup> <https://techcrunch.com/2008/04/17/microsoft-acquires-farecast-for-115m/>

Accenture che lavora con diverse aziende installando sistemi wireless per esempio su un gran numero di autobus. In questo modo l'impresa monitora i motori in modo da prevenire guasti e riducendo i costi della manutenzione (ridotti di un rilevante 10%)<sup>9</sup>.

- **Analisti di big data:** sono comprese all'interno di questa componente tutte quegli individui e società che sono caratterizzati da un forte orientamento culturale verso i big data. Un esempio chiarificatore è la figura di Bradford Cross, fondatore di FlightCaster.com e Prismatic. Il primo prevedeva la probabilità di ritardo di un volo interno agli Stati Uniti, analizzando tutti i voli degli anni precedenti e messi in correlazione tra loro. La seconda, operante nel settore dei media dedicati all'informazione, associava e classificava contenuti, sulla base dell'analisi del testo, contenuti ricavati da tutto il web;

il sistema non distingueva se la fonte della pubblicazione e se il contenuto si valutava rilevante e popolare, sulla base dei click e condivisioni, sarebbe apparso in cima alla classifica.

Cross è riuscito a vedere ciò che è possibile fare, con le informazioni pubbliche, cosa che invece non sono stati in grado di fare i possessori dei dati.

Vi sono imprese quali Google ed Amazon che raccolgono invece tutte e tre le suddette categorie ovvero hanno i dati, l'expertise e l'orientamento culturale, ma differiscono tra loro per le strategie utilizzate. Se infatti Google quando raccoglie i dati ha in mente degli utilizzi secondari, Amazon invece è più concentrata sul primario utilizzo che questi possono avere e solo parzialmente verso un utilizzo secondario che le informazioni generate possano avere.

---

<sup>9</sup> Mayer Schönberger V., Cuckier K, *Big Data: una rivoluzione che trasformerà il nostro modo di vivere e già minaccia la nostra libertà*. Milano: Garzanti, 2013

Google ad esempio quando ha impiegato le sue auto per *Street view* non si limitava a raccogliere dati per le mappe, ma anche per sperimentare i sistemi di guida automatica.

Amazon invece considera i dati relativi alle ricerche di prodotti, ma non li ha mai utilizzati per scopi extra ordinari. Sempre Amazon non dà l'impressione di utilizzare i dati raccolti attraverso gli e-book Kindle, i quali sono in grado di rilevare note e sottolineature da parte degli utenti ma non vende tali informazioni né agli autori né agli editori, i quali potrebbero sfruttare tali informazioni in vari modi possibili e vendere così i libri in modo più funzionale. Da ciò si può notare come i big data produca una trasformazione nei modelli di business utilizzati dalle imprese ed inoltre anche i rapporti di interazione tra partner.

### 1.3.1 IL FUTURO DEI BIG DATA

Ipotizzando quello che potrebbe essere il futuro dei big data, è possibile introdurre la curva ad "S" del cambiamento tecnologico. Essa ha come variabili: sull'asse delle ascisse l'impegno considerato come quantitativo degli investimenti e il grado di impegno a livello organizzativo destinati allo sviluppo di una certa tecnologia, nel nostro caso i big data; sull'asse delle ordinate troviamo invece il livello di performance inteso come valore ottenuto.

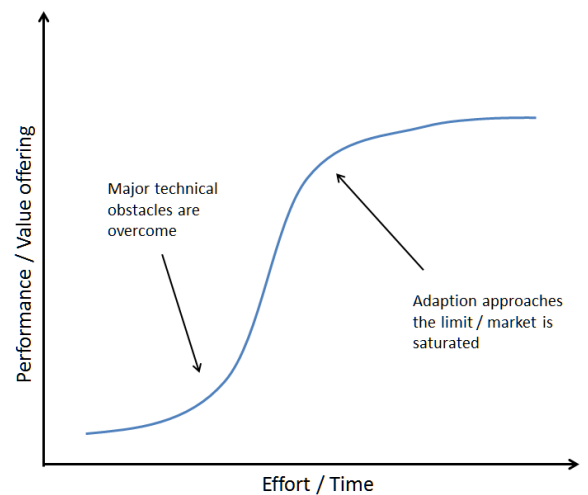


Figura 6: Curva ad S del miglioramento tecnologico



“Si ritiene che la quantità di dati digitali possa avere un andamento simile alla curva tecnologica a S che descrive il limite generico che ha una tecnologia rispetto all’aumentare del suo impiego<sup>10</sup>”.

Si ritiene infatti plausibile associare il futuro andamento dei big data, con la curva ad “S” della tecnologia, poiché nel nostro caso appunto (proprio come ipotizzato in precedenza con la legge di Moore), la capacità di immagazzinare i dati dei database elettronici cresce esponenzialmente come anche a crescere in maniera più che proporzionale sono il numero degli abbonamenti sottoscritti di sim mobile (vedi figura 6).

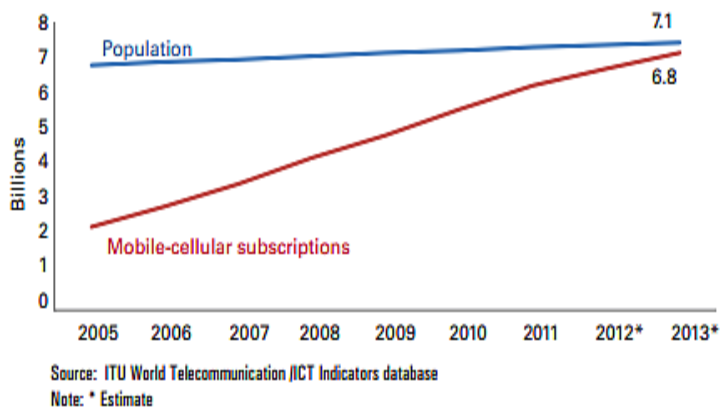


FIGURA 7: NUMERO DEGLI ABBONATI A CONTRATTI PER CELLULARI

Ritornando al grafico proposto in figura 5, possiamo vedere come in una prima fase, risulta lento il miglioramento della performance causato dal fatto che gli strumenti che si occupano della raccolta e dell’analisi dei dati sono ancora poco sviluppati e soprattutto il costo risulta piuttosto elevato. In questa fase infatti ancora non è stato trovato un metodo standardizzato di elaborazione dei dati e molte delle risorse a disposizione vengono utilizzate per valutare soluzioni alternative.

<sup>10</sup> Giuseppe Roccasalva, 201, p. 97, “I Big Data e gli strumenti di visualizzazione analitica: interazioni e studi induttivi per le P.A.”.

Nella seconda fase invece, al crescere dell'impiego di fattori produttivi e del volume degli investimenti, verranno acquisite conoscenze maggiori e il miglioramento tecnologico comincerà a crescere in maniera più che proporzionale. Arrivati infine ad un certo momento della crescita, il rendimento della tecnologia come anche le risorse impiegate per lo sviluppo di questa inizieranno a decrescere. La tecnologia tenderà quindi al proprio limite naturale fino a quando il costo marginale di ogni miglioramento risulterà maggiore al beneficio procurato dalla tecnologia stessa<sup>11</sup>.

Come abbiamo visto nel paragrafo 1.1, nel grafico “*How much information is there in the information society?*” il numero dei dati negli ultimi anni digitalizzati è cresciuto in maniera più che proporzionale, quindi con ogni probabilità possiamo definire che ci troviamo proprio nella fase iniziale della curva ad S del miglioramento tecnologico.

### 1.3.2 GLI INTERMEDIARI DEI DATI

Oggi vediamo come la maggior attenzione e quindi la leadership, nella catena del valore dei big data, sia posta su coloro che hanno il giusto orientamento culturale e idee innovative rispetto all'utilizzo dei dati. Con il tempo però, probabilmente altre aziende adotteranno il giusto orientamento verso i dati, diventando così non più un punto di forza, bensì un valore standard, in questo modo il vantaggio dei pionieri si ridurrà in termini relativi. Quindi , probabilmente saranno i dati stessi ad essere il maggior valore, proprio come in una miniera d'oro, quando gli strumenti di estrazione saranno acquisibili facilmente ed a basso prezzo, sarà la materia prima stessa, nel nostro caso, nei dati risiederà il massimo valore.

---

<sup>11</sup> Fontana F., Caroli M.: *Economia e Gestione delle Imprese*, McGraw-Hill, (2013). Cap 11.

Le imprese, o comunque coloro che detengono i dati, comprenderanno il potenziale valore di questi e saranno più restii a cederli o comunque se lo dovessero fare lo faranno a caro prezzo. In questo scenario si verrà ad identificare la nuova figura degli “intermediari di dati”, i quali avranno le capacità di raccogliere da fonti differenziate, aggregare ed impiegare i dati per scopi innovativi; i possessori dei dati acconsentiranno allo scambio di dati con questi enti, poiché in alcune situazioni il valore si potrà estrarre solo attraverso l’intermediazione di questi.

Inrix ad esempio, società specializzata nell’analisi del traffico automobilistico, si occupa di aggregare dati in tempo reale attraverso geolocalizzazione e sistemi gps, inviati da milioni di automezzi: taxi, telefoni cellulari, ma soprattutto Inrix detiene accordi con aziende automobilistiche quali bmw, ford e toyota. In particolare essa combina i dati ottenuti dagli automezzi e dagli automobilisti, con le informazioni storiche relative all’andamento del traffico e alle condizioni atmosferiche, in modo da poter offrire in tempo reale aggiornamenti ai propri utenti<sup>12</sup>

Le case automobilistiche hanno tutto l’interesse che, a fare questo lavoro, siano gli intermediari, le quali nonostante possiedano milioni di dati, agendo autonomamente, non sarebbero in grado di assicurare previsioni accurate né complete.

Oggi i dati costituiscono una materia prima del mercato ed un asset che può avere più funzioni rispetto all’originale di partenza. Ad esempio i dati sul traffico raccolti dall’azienda Inrix, sono stati utilizzati per analizzare il benessere delle economie locali. L’azienda ha venduto i dati relativi al traffico ad un fondo di investimenti, il quale ha utilizzato le informazioni sul traffico intorno ai punti vendita di una catena di supermercati per stimare le vendite;

---

<sup>12</sup> Mayer Schönberger V., Cuckier K, *Big Data: una rivoluzione che trasformerà il nostro modo di vivere e già minaccia la nostra libertà*. Milano: Garzanti, 2013

si è scoperta infatti una correlazione positiva tra l'aumento del traffico automobilistico e l'incremento delle vendite. Il valore quindi si è trasferito dalle competenze all'idea e si sta spostando oggi verso le informazioni e compito difficile per gli intermediari sarà quello di convincere le aziende che condividere le informazioni porterà un vantaggio anche a queste. A riprova del fatto che i dati stiano acquisendo una posizione di leadership è dimostrata dalla differenza di prezzo d'acquisto tra le due aziende, infatti ITA Software, (come già analizzato nel paragrafo 1.3.1) è stata pagata circa sette volte il valore di Farecast, per il fatto che, a differenza di quest'ultima, fosse detentrica di dati.

- 2006, Microsoft acquista Farecast per \$115 milioni<sup>13</sup>
- 2008, Google acquista ITA Software per \$700 milioni.

Con l'avvento dei big data si è introdotta la figura professionale degli intermediari di dati. Tale avvento porterà probabilmente ad una perdita d'influenza degli esperti in moltissimi settori. Nei massmedia i contenuti pubblicitari sono sistematicamente stabiliti dai dati. Questi infatti sono in grado di suggerire meglio gli argomenti che interessano al lettore, rispetto a quanto possano fare i migliori giornalisti. Questo dimostra che stanno mutando le competenze necessarie per competere nell'ambiente del lavoro. Infatti i concorsi organizzati dalla piattaforma online per progetti sui big data Kaggle, mostrano come i precursori dei big data provengano da ambienti diversi rispetto a quelli in cui incidono e sono quindi sempre estranei al settore in cui riescono a generare risultati di successo.

Ancora, nel settore video ludico, quello che prima si basava sulla creatività dei produttori, (trama, personaggi, ritmo di gioco ecc...) è oggi affidata in buona parte all'analisi dei dati. La società Zynga, produttrice di una serie di

---

<sup>13</sup> Vedi 1.3.1 "La catena del valore dei big data" nota 8.

videogames, (FarmVille, FishVille ed altri) legati al contesto del social network Facebook, ha la possibilità di rilevare i problemi e porvi rimedio ed inoltre modificare il gioco sulla base dei gusti e delle preferenze dell'utente. Nello specifico, gli analisti della società, analizzano i data rilevati per stabilire connessioni tra le vendite e altri fattori, dai colori o dal desiderio di imitare i loro amici.

La struttura di molte categorie verrà riprogettata, infatti in un mondo in cui i dati divengono fattori competitivi, la competitività non dipenderà più dalle dimensioni fisiche delle infrastrutture, bensì dalle dimensioni degli archivi e dalla velocità di processazione dei dati. In questo scenario anche le piccole imprese potranno divenire competitive a patto che detengano una forte presenza virtuale. Potrebbe emergere anche una nuova categoria di detentori dei dati ovvero, gli individui; le persone infatti potrebbero reclamare i propri diritti sulle informazioni e questo potrebbe dar loro un maggiore potere, considerando anche l'idea che un giorno sia il singolo cittadino a stabilire di vendere e commercializzare i propri dati personali.

#### 1.4 L'ALTRA FACCIA DELLA MEDAGLIA

I big data risultano quindi essere uno strumento molto utile, per le grandi società, per i governi come anche per gli scienziati, consentono di stabilire correlazioni tra eventi e situazioni, che mai prima di ora, e siamo solo all'inizio, avremmo pensato avrebbero potuto legare tra loro. Bisogna però fare attenzione però che la fiducia nei dati non sfoci in una ossessione per i dati, giustificata, apparentemente, dalla loro "onnipotenza".

Nel 2008 Google sviluppò una piattaforma che era in grado di anticipare la diffusione dell'influenza negli Stati Uniti, analizzando le ricerche effettuate dagli utilizzatori di internet monitorando le ricerche relative alla malattia, confrontando tali informazioni con l'elenco di dati forniti dal CDC (centers for Disease Control and Prevention) relative al periodo 2003-2008. Gli ingegneri di Google processando e testando le queries hanno trovato una congiuntura di 45 parole chiave che, se impiegate insieme in un modello matematico mostravano una forte correlazione tra le loro previsioni e i dati relativi alla locazione dell'influenza<sup>14</sup>

Google è riuscita così ha dire, in tempo reale dove si era propagata l'influenza. Questo metodo è risultato uno strumento sia molto efficace e meno dispendioso, dal punto di vista economico, sia anche privo di fondamenti teorici, il team di Google infatti si limitò a raccogliere milioni di ricerche e gli algoritmi sviluppati *ad hoc* si occuparono del resto. La piattaforma sviluppata dal colosso di *Mountain View* ebbe notevole successo poiché: era in grado di fornire un'analisi accurata dei dati (ogni informazione era infatti sottoposta a setaccio), perdeva di significato la comprensione del rapporto di causa effetto, lasciando posto alla correlazione e lasciando da parte modelli statistici quali per esempio il campionamento; infine citando il saggio provocatorio pubblicato su "Wired", "*The end of the teory*": con sufficienti dati, i numeri parlano da soli.

Nel 2012-2013 però il Google Flue Trends, dopo aver fornito per anni risultati attendibili riguardo il diffondersi dell'influenza negli Stati Uniti, il modello sviluppato da Google perdeva le sue capacità di analisi. Il modello infatti ricco di dati ma privo di analisi teoriche aveva sovrastimato i risultati dell'epidemia di circa il 40%, portando il modello ad un fallimento.

Uno dei motivi per cui Google ha fallito la sua analisi è stato il fatto che si era creato un forte allarmismo intorno all'influenza sviluppatasi nel 2012,

---

<sup>14</sup> A.F. Dugas et al., Google Flu Trends: *Correlarion with Emergency Department Influenza Rates and Crowding Metrics*; CID Advanced Access, 8 gennaio 2012

questo ha portato una quantità innumerevole di ricerche rispetto ai sintomi della malattia, anche di cittadini che non li avevano riscontrati su loro stessi, ma che lo facevano semplicemente per acquisirne informazioni.

Un'analisi che esclude qualsiasi fondamento teorico e che si basa esclusivamente su correlazioni tra i dati, risulta sicuramente, fragile. Se infatti non si è a conoscenza di cosa vi sia dietro le correlazioni, questo porterà a non capire le cause del malfunzionamento dell'analisi di correlazione.

Google Flue Trends si riprenderà attraverso l'utilizzo di nuovi dati e ricalibrando gli algoritmi utilizzati per l'analisi. Ci sono comunque molte ragioni per essere entusiasti dei dati, dal momento che ci offrono nuove possibilità più efficaci di raccogliere ed analizzare enormi masse di dati, occorrerà però imparare dagli errori commessi da Google.

Gli statistici hanno passato gli ultimi 200 anni per comprendere quelle che sono le trappole, quando cerchiamo di interpretare il mondo attraverso i dati. Oggi i dati sono più grandi, più veloci e soprattutto più economici, non dobbiamo però fingere che le trappole siano state tutte messe in sicurezza<sup>15</sup>. Ogni medaglia ha una seconda faccia, anche i big data, infatti sia i numeri sia i dati risultano strumenti cruciali in molte situazioni per comprendere situazioni complesse, ma decontestualizzati (come appreso dal caso Google Flue Trends) i numeri restano numeri e alcune volte divenire fuorvianti.

Un'altra problematica che può essere riscontrata, relativamente ai big data è quella della privacy. Questi consentono infatti un maggior controllo sulle nostre vite e in questo modo rendono inadatti e obsoleti gli strumenti finora a tutela della nostra riservatezza. Con l'avvento dei big data infatti si modifica l'essenza del rischio per la privacy, infatti il valore dei dati non

---

<sup>15</sup> Harford T., Financial Times, "Big data are we making a big mistake?", 28 marzo 2014.

risiede più solamente nel loro scopo primario, ma soprattutto nei loro utilizzi secondari. Nell'era dei big data: consenso informato, dissociazione e anonimizzazione, tecniche di base per la salvaguardia della privacy, hanno perso gran parte della loro efficacia. Le aziende che vogliono raccogliere le nostre informazioni al fine di riutilizzarle, devono ricevere il “nostro” consenso, questo però diviene più complicato in un mondo dominato dai big data, poiché alcuni degli utilizzi secondari perseguibili oggi, non erano stati neppure immaginati (durante la raccolta dati). Per quanto riguarda l'anonimizzazione dei dati: depurare la massa dei dati da tutti gli elementi che possano permettere un'identificazione della persona, questo sistema risulta funzionare solo in un mondo caratterizzato da small data. Nel 2006 il New York Times analizzando una serie di queries raccolte dalla società AOL, che erano state effettivamente anonimizzate, è riuscita a risalire a nome e cognome, residenza di molti dei soggetti “anonimizzati”<sup>16</sup>. In questo caso i dati personali venivano rivelati dal contenuto delle ricerche degli utenti. Con ogni probabilità i big data porteranno benefici alla società del domani, ma allo stesso tempo ci porteranno ad affrontare nuove problematiche con nuove soluzioni.

---

<sup>16</sup> ARRINGTON M., TECHCRUNCH: *FIRST PERSON IDENTIFIED FROM AOL DATA: THELMA ARNOLD*, 2006.



## 2. APPLICAZIONE DEL FENOMENO DEI BIG DATA

### 2.1 IL FUTURO MERCATO DEI BIG DATA

A detta del presidente e ceo di Dell il mercato dei big data sarà un mercato da bilioni di dollari. IDC, azienda americana specializzata in ricerche di mercato, information technology ha formalizzato ipotesi più modeste nonostante siano comunque positive; IDC ritiene infatti che il mercato dei big data possa crescere ad un tasso del 23% annuo, raggiungendo nel 2019 48.6 bilioni di dollari<sup>17</sup>. Ci si aspetta inoltre, che le opportunità di servizio correlate, rappresentino la metà dei ricavi delle previsioni. Dopo i servizi, i software saranno i grandi generatori di profitto, sempre secondo IDC infatti si prevedono ricavi pari a 55 miliardi di dollari nel 2019, in particolare quasi la metà si prevede che proverranno da acquisti di ricerca per l'utente finale, strumenti di analisi e strumenti di gestione di data warehouse<sup>18</sup>.

---

<sup>17</sup> DAVIS J., INFORMATIONWEEK: BIG DATA, ANALYTICS SALES WILL REACH \$187 BILLION BY 2019, MAGGIO 2016

<sup>18</sup> Press G., Forbes: 6 Predictions For Big Data Analytics And Cognitive Computing In 2015, ottobre 2015

Secondo le previsioni questi numeri sono destinati ad incrementare, sempre secondo IDC i dati che saranno elaborati all'interno di un'azienda nel 2020 saranno il doppio di quelli di oggi.

Nel 2014 solo il 10% delle imprese ha messo sul mercato i dati raccolti, mentre solo un anno dopo la percentuale delle imprese è divenuto il 30%, generando un aumento del 200%.

La Forrester Research ha dichiarato che oggi tutte le imprese si trovano nel “*data business*” e che però la maggior parte di queste non sono ancora in grado di gestire in maniera ottimale i dati raccolti, ma dovranno farlo per poter prendere migliori decisioni in *real time*.

Alcuni dati interessanti sono stati riportati da un recente articolo: “*Big Data and Advanced Analytics Survey 2015*” volume 1 della Evans Data Corporations. Si tratta di un'indagine che si propone di offrire una visione strategica degli atteggiamenti e dei modelli di adozione di coloro che si relazionano con i Big Data. In particolare riporto alcune delle analisi più interessanti:

1. Le informazioni raccolte più frequenti all'interno delle banche dati risultano essere: per il 9,6% vendite e informazioni sui clienti, per il 9,4% analisi degli IT e per l'8,4% le transazioni finanziarie. (Vedi fig.8)

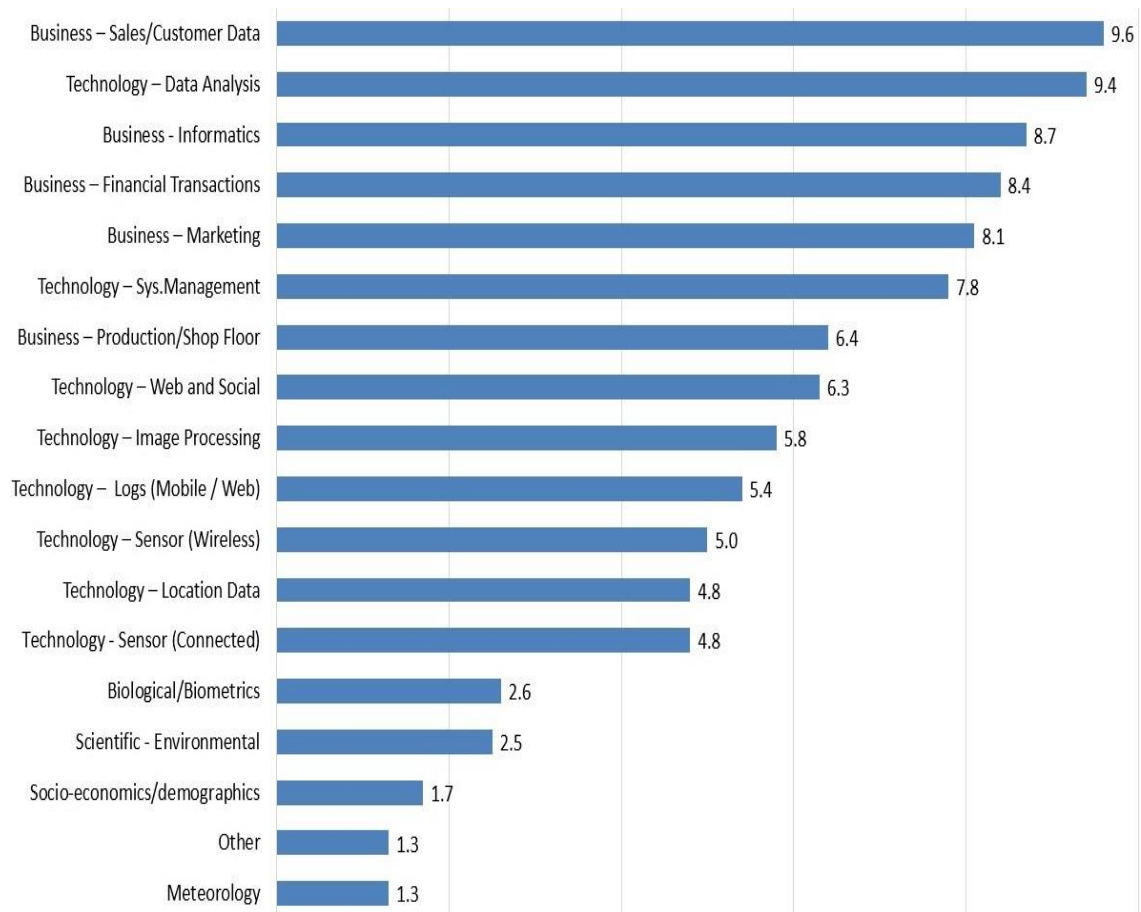


Figura 8: "What kind of information feed your company's data stores?"

- Le divisioni, che all'interno delle organizzazioni scansionate dalla Evans Data, risultano le maggiori utilizzatrici di dati e strumenti analitici per l'analisi di questi sono: la divisione marketing, che utilizza il 14,4% dei big data, l'IT il 13,3% ed il 13,0% è utilizzato dal settore ricerca e sviluppo.

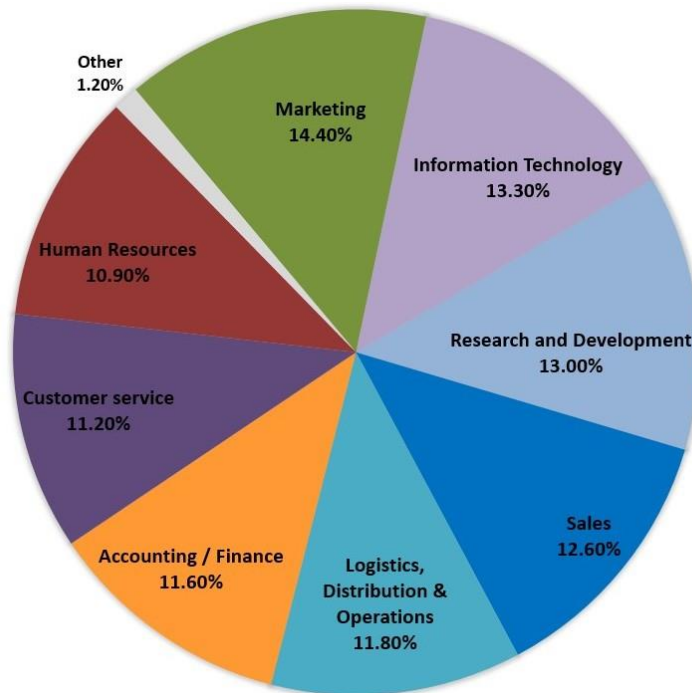


Figura 9: Which department in your organization are using advanced data analytics or Big Data solutions for your company's needs?"

### 2.1.1 ANALISI SWOT

I big data risultano essere una risorsa primaria per poter competere in un mondo sempre più digitalizzato. Capirne i punti di forza e i limiti, sarà determinante per ricavarne il massimo beneficio da questi. A tale scopo sarà utile mettere in luce, attraverso la SWOT analysis, punti di forza, punti di debolezza, opportunità e minacce.

- Punti di forza: il maggior punto di forza del modello risiede nelle 4V: volume, varietà, velocità di trasmissione dei dati ed il valore generato dai dati, queste stanno infatti aprendo una serie vastissima di opportunità in tutte le aree funzionali di business: nel marketing, nella contabilità, nella finanza e nella gestione delle risorse umane. I big data, come visto nei paragrafi precedenti,

sono destinati ad una crescita via via sempre maggiore in tutte le quattro aree “V”, grazie al maggior utilizzo di social network, all’incremento degli acquisti online e al “Internet of Things”. Questo costituisce un punto di forza naturale. Infine un terzo punto di forza è costituito dalla maggiore flessibilità tecnica data dal fatto che le analisi che nel passato avevano a disposizione solo un limitato campione di dati, oggi invece grazie all’enorme massa di dati il campione non risulta più limitato.

- **Punti di debolezza:** in primo luogo dagli studi svolti risulta che negli Stati Uniti manchino persone che abbiano capacità di analisi profonda ed inoltre managers che siano in grado di capire i dati, fin quando questa mancanza non verrà risolta, questo risulterà un punto di debolezza grave. Un secondo punto di debolezza è rappresentato dalla necessità di possedere infrastrutture tecnologiche che siano capaci di acquisire, memorizzare e analizzare l’ingente quantità di dati, senza una struttura adatta infatti, le conoscenze offerte dai big data saranno sempre limitate. Con la grande quantità di dati a disposizione però gli analisti, soprattutto se con limitata capacità di analisi, potrebbero commettere false associazioni tra i dati che porterebbero a gravi conseguenze gestionali.
- **Opportunità:** i big data hanno una serie molto ampia di benefici in diversi settori, nella produzione, nella catena di distribuzione e nei trasporti e come abbiamo visto in precedenza si estende a più aree funzionali quali il marketing, la contabilità, la finanza e la gestione delle risorse umane. I big data possono essere utilizzati per estrarre interpretare i desideri dei clienti, in modo tale che i prodotti offerti riescano a soddisfare sempre le loro esigenze.
- **Minacce:** all’aumentare dei dati a disposizione, aumenteranno anche le informazioni personali che saranno a disposizione delle aziende e di chiunque possa passare in rassegna i dati, si viene quindi ad identificare un problema relativo alla privacy. Inoltre l’aumento dei dati comporta la necessità di spendere risorse per “pulire” i dati prima di poterli utilizzare nelle proprie analisi.

### PUNTI DI FORZA

- LE 4 V DEI BIG DATA
- ESPANSIONE DELLA PORTATA E DELLA PROFONDITA' DEI DATI
- MAGGIORE FLESSIBILITA' TECNICA

### PUNTI DI DEBOLEZZA

- ASSENZA DI PROFESSIONISTI QUALIFICATI
- INADEGUATEZZA DELLE STRUTTURE
- ASSENZA DI UNA PROSPETTIVA CHIARA

### OPPORTUNITA'

- PRODOTTI E SERVIZI PIU' «SMART»
- MARKETING PIU' EFFICACE
- RIDUZIONE DELLE FRODI
- PERSONALIZZAZIONE

### MINACCE

- RIDUZIONE DELLA PRIVACY
- CORRELAZIONI ERRONEE DI DATI

## 2.2 DATA-DRIVEN BUSINESS MODEL

Nel primo paragrafo abbiamo visto che le imprese che utilizzano i dati come risorsa primaria, siano in continuo aumento. Le imprese devono quindi adattarsi al cambiamento prodotto dall'arrivo della nuova risorsa: i (big) data.

Esploriamo ora quelli che potrebbero essere i modelli di business che utilizzano come risorsa chiave di maggiore importanza i (big) data.

Sebbene il modello di business che si basa sui dati (*Data-driven business model*) non abbia ancora una definizione specifica, potremmo indicarlo come quel modello che fa affidamento sui dati, quale risorsa chiave. Da questa definizione possiamo ricavare tre implicazioni:

1. Il modello non è limitato a quelle aziende che conducono analisi analitiche, ma include anche quelle che si occupano semplicemente di raccogliere e aggregare i dati.
2. Include non solo le aziende che vendono dati o informazioni, ma anche quelle che vendono qualsiasi altro prodotto o servizio che si basano sui dati quale risorsa centrale.
3. Ogni impresa utilizza i dati in qualche maniera all'interno del proprio business, il modello analizzato include quelle aziende che usano i dati quali risorsa fondamentale del proprio business.

Il modello DDBM (*Data-driven business model*) consiste in sei dimensioni comuni alla maggior parte dei business models che sono<sup>19</sup>:

1. Le risorse chiavi: “tutte le attività, le capacità, i processi organizzativi, gli attributi di un'impresa, le informazioni e la conoscenza in possesso di un'azienda.” (Barney,1991). Dalla definizione proposta si possono identificare i dati come risorsa chiave del DDBM. Sono state individuate sette fonti di dati, che possono essere suddivise in interne ed esterne. Tra quelle interne ritroviamo: i dati che sono in possesso dell'impresa, quelli che possono essere generati dai sistemi informatici interni, quelli che possono essere raccolti attraverso analisi sul web o attraverso dispositivi di monitoraggio fisico. I dati esterni invece comprendono i dati acquisiti, che

---

<sup>19</sup> Hartmann P., Zaki M., Fieldmann N., Neely A., University of Cambridge: *Big Business? A taxonomy of Data-driven Business Models used by Start-up firms*, marzo 2014.

possono essere acquistati da fornitori di dati e società quali Twitter, Facebook ecc., poi i dati che possono essere forniti da clienti, dati che sono disponibili al pubblico senza costi aggiuntivi (open data).

2. Le attività chiave: nelle tradizionali imprese che sono orientate alla produzione si utilizza la catena del valore realizzata da Porter per rappresentare le attività che generano valore. Questo modello però poco è poco adeguato per rappresentare modelli di business che hanno come fonte di valore i dati. È stata quindi sviluppata una catena del valore “virtuale”<sup>20</sup> lungo la quale si sviluppano cinque attività chiave che sono: la raccolta, organizzazione, selezione, sintesi e distribuzione dei dati.
3. Il valore offerto: è la dimensione centrale di qualsiasi business model e può essere definito come l’espressione dell’esperienza che un cliente riceverà dalla proposta del fornitore. Nel caso delle imprese caratterizzate dal DDBM, il valore offerto può essere diviso in due categorie, che sono i dati grezzi e la conoscenza ai quali se ne può aggiungere un terzo che corrisponde a qualsiasi offerta non virtuale.
4. Il segmento di clientela: questa dimensione ha a che fare con il target dell’offerta. Il metodo generalmente più utilizzato per segmentare la clientela è quello di dividere la clientela in businesses (B2B) e consumatori (B2C).
5. Il modello di reddito: le aziende per sopravvivere nel tempo hanno bisogno di almeno un flusso di entrate. Si possono identificare sette flussi di entrata quali: vendita di attività, prestito-noleggio-locazione, le licenze, il diritto di utilizzo di un bene, un canone di abbonamento; la pubblicità o infine una tassa di brokeraggio per un servizio di intermediazione.

---

<sup>20</sup> Rayport J., Sviokla J., HBR: *Exploiting the virtual chain value*, novembre-dicembre 1995.



6. La struttura dei costi: per generare reddito le imprese devono sostenere una serie di costi per acquistare materie prime, tecnologie, prodotti ecc. Nel caso di aziende che utilizzano i dati per generare valore, esse avranno uno specifico vantaggio di costo, se i dati utilizzati nei suoi prodotti o servizi, sono stati generati indipendentemente dall'offerta specifica. Un esempio è Twitter, il quale può utilizzare le informazioni raccolte dai *tweet*, senza costi aggiuntivi, per poter svolgere un servizio di analisi.

È stato riscontrato che la produttività delle aziende che hanno come business model il “Data-driven” hanno una produttività maggiore del 5-6% rispetto ad aziende simili, che invece non utilizzano lo stesso processo<sup>21</sup>. Esiste chiaramente un effetto positivo scaturente dall'utilizzo dei big data all'interno delle organizzazioni, molto dipende dalla capacità delle aziende di riuscire a capitalizzare tali vantaggi offerti dai big data.

---

<sup>21</sup> E. Brynjolfsson; L. Mitt; H. Kim, Social science research network paper: “Strength in Numbers: *How does Data-driven Decision Making Affect Firm Performance?*” (2011)

## 2.2.1 QUADRO TEORICO PER LA CREAZIONE DI UN DBBM

È fondamentale quindi che le aziende che vogliono sfruttare al meglio il vantaggio proposto dai dati, abbiano una struttura adeguata che permetta loro di capitalizzare tale vantaggio.

Secondo il blueprint pubblicato dalla University of Cambridge: “Data and Analytics – Data – Driven Business Models” (2015), per creare un business model che sia orientato ai dati bisogna rispondere a 6 domande fondamentali:

### 1. Cosa vogliamo ottenere dall'utilizzo dei big data?

Stabilendo fin da subito il risultato desiderato il business può mantenere il suo focus su un obiettivo che sia realistico in modo da ridurre lo spreco di risorse umane e di tempo durante il processo di ricerca e analisi. Dall'analisi svolta sulle organizzazioni selezionate risultano sette vantaggi competitivi chiave: una catena del valore abbreviata, l'espansione; il consolidamento dell'organizzazione; la velocità di elaborazione dei dati; la differenziazione ed il brand.

Per esempio la famosa catena di abbigliamento Zara aveva lo scopo di accorciare la propria catena del valore in modo tale da poter promuovere la propria offerta di prodotti sul mercato, nel più breve tempo possibile. Inserendo statistiche di vendita in tempo reale e analizzando nei propri sistemi analitici i post sui social media, Zara è in grado di accaparrarsi le emergenti tendenze di mercato.

Ne è un esempio pratico, il caso del vestito indossato dalla musicista Beyoncé nella notte di apertura del suo tour mondiale. Zara analizzando grandi volumi di dati non strutturati, prodotti sui social media, è riuscita a mettere sul mercato un prodotto ispirato a quello della cantante prima ancora che il tour

arrivasse a conclusione, in questo modo Zara è riuscita a soddisfare in tempo reale la domanda emergente.

2. In cosa consiste il valore offerto?

Le imprese possono offrire sia i dati grezzi, intesi come semplici dati, senza alcun significato ulteriore, ma possono anche offrire dati che sono stati sottoposti ad un lavoro interpretativo e che quindi possono essere definiti come informazioni/conoscenza.

3. Di quali dati abbiamo bisogno e come faremo per acquisirli?

I dati sono ovviamente l'elemento fondamentale per la costruzione di questo modello è inoltre importante per il successo dello stesso, stabilire quali dati siano i più adatti. I business che hanno un considerevole numero di clienti e di conseguenza un elevato potenziale di interazioni con i clienti stessi, sono ben posizionati per utilizzare in modo efficace i dati che spesso sono combinati con dati provenienti da altre fonti (open data o dati che possono essere acquistati).

Ad esempio, la famosa catena di abbigliamento Topshop combina i dati acquisiti dai clienti, gli open data generati da blog e social network con i dati posseduti all'interno dei suoi database per poter determinare le nuove tendenze, in un settore altamente competitivo quale quello dell'abbigliamento. Senza questi processi di analisi dei dati interni ed esterni, i rivenditori perderebbero una significativa opportunità di guadagno.

4. In che modo i dati saranno elaborati ed applicati?

L'elaborazione dei dati rivela il vero valore contenuto all'interno di essi. È importante conoscere quali attività chiave saranno utilizzate nell'elaborazione dei dati, in modo da poter pianificare l'organizzazione. Come abbiamo visto in precedenza è stata sviluppata una catena del valore virtuale, che permette di inquadrare quelle che sono le attività fondamentali.

Le società possono raccogliere dati sia generandoli internamente sia acquisendoli da fonti esterne. L'elaborazione è svolta attraverso attività di analisi che può essere suddivisa in: analisi descrittive, attività di analisi che spiegano il passato, attività di analisi di previsione; analisi che predicono un esito futuro e ne suggeriscono decisioni.

5. In che modo diventano fonte di reddito?

Sono state identificati sette flussi di reddito (vedi paragrafo precedente 6.) Il modello di entrate più utilizzato dalle organizzazioni prese in analisi risulta essere la pubblicità. IL ceo del New York Times ad esempio, resosi conto che il numero di copie fisiche vendute si era ridotto, portando ad un calo dei ricavi, ha deciso di offrire ai suoi utenti online un servizio di lettura gratuita. Senza costi d'accesso il Times è riuscito ad accumulare un elevato numero di click dai suoi utenti, riuscendo a costruire un profilo unico per ogni lettore, in questo modo ha garantito alle pubblicità di effettuare pubblicità mirate su ogni cliente.

6. Quali sono le barriere che ci impediscono di raggiungere il risultato desiderato?

È importante che il personale sia orientato ai dati e che inoltre abbia un'esperienza sufficiente all'interno di tale business in modo tale che non si sviluppi una percezione negativa del DDBM. Seminari di formazione in cui vengono esposti i benefici apportati, possono ridurre la resistenza del personale nell'aprirsi alle nuove procedure basate sui dati.

## 2.3 NUOVI RUOLI NECESSARI GESTIRE UN BIG DATA TEAM

Le aziende sembrano prestare troppa attenzione a come trasformare l'insieme dei dati in valore e troppa poca sul fattore umano della questione.

Valorizzare il fattore umano, non consiste solamente nell'assumere persone che siano specializzate ma anche e soprattutto considerare due elementi fondamentali:

1. Identificare i ruoli necessari all'interno dell'azienda
2. Costruire una mentalità che sia orientata al “servizio clienti” nell'ufficio di analisi avanzate<sup>22</sup>.

Oggi infatti risulta difficile intercettare quelli che sono gli specialisti nel settore dei big data. Si prospetta che nel 2018 negli Stati Uniti ci sarà una carenza tra i 140,000 e i 190,000 dipendenti con competenze di analisi profonde e di circa 1,5 milioni di managers e analisti che abbiano il know-how per estrapolare dagli innumerevoli dati, le informazioni per poter prendere le giuste decisioni<sup>23</sup>.

Possiamo quindi identificare cinque ruoli importanti per le imprese che si trovano a dover gestire un elevato numero di informazioni.

1. **Data hygienists**: aumentando il numero delle informazioni, di sicuro aumenta anche il numero di quelle reputate irrilevanti. Il “data hygienist” svolge il suo lavoro prima dell'analista dei dati, dell'architetto di database ecc. Questo ruolo funge da supporto al Project Manager che ha bisogno di riportare in tempo reale informazioni reali ai superiori. Questo ruolo si

---

<sup>22</sup> Aricker M., McGuire T., Perry J., Harvard Business Review: *Five Roles You Need on Your Big Data Team*, luglio 2013.

<sup>23</sup> Manyika J., Chui M., Brown B., McKinsey Global Institute: *Big Data: The next frontier for innovation, competition, and productivity*, maggio 2011.

occupa quindi di raccogliere, depurare ed infine convalidare i dati ancora prima che interagiscano con l'azienda.

2. ***Data Explorers***: si occupano di passare in rassegna enormi masse di dati per trovare quelle di cui l'azienda ha realmente bisogno. Questo rappresenta un ruolo molto importante, poiché le informazioni che sono utili, difficilmente si trovano organizzate e raccolte come in modo tale da essere di facile accesso. Un esempio di facile comprensione è il registratore di cassa il quale scopo iniziale era quello di tenere traccia degli acquisti dei clienti, tali dati però raccolti e organizzati possono essere utilizzati per prevedere i futuri acquisti del cliente.
  
3. ***Business Solution Architects***: mettono insieme i dati raccolti e li organizzano in modo tale che siano pronti per essere analizzati. I dati vengono strutturati in modo tale che possano essere utilmente sfruttati in ogni momento da tutti gli utenti. Spesso i dati infatti, per essere utili, hanno bisogno di essere aggiornati ogni ora e a volte anche ogni minuto.
  
4. ***Data scientist***: è stata definita dall'economista Ronald Varian come la professione più interessante del futuro. Tale figura deve avere più competenze: la prima è quella di saper acquisire, gestire, organizzare ed elaborare i dati. La seconda è quella di tipo statistico, saper come e quali dati estrarre infine deve saper comunicare all'interno dell'azienda cosa suggeriscono i dati. La caratteristica principale di tale figura deve essere la curiosità, intesa come attitudine ad analizzare in profondità i fenomeni senza fermarsi all'apparenze ed identificare una serie di ipotesi da verificare con l'aiuto dell'analisi e lo studio dei dati. Il "data scientist" dovrà essere dotato

di una solida conoscenza del campo dell'informatica, saper comprendere gli aspetti tecnologici e conoscere quelle che sono le problematiche e le sfide all'interno del settore di interesse. Egli quindi ha il compito di creare, attraverso dati organizzati, di creare modelli analitici, che permettano di fare previsioni sul comportamento dei clienti e segmentare il mercato.

5. *Campaign Experts*: hanno l'incarico di trasformare i modelli in risultati, hanno conoscenze specifiche dei sistemi per produrre campagne marketing specifiche, del tipo a quali clienti e quando mandare un certo messaggio. Utilizzano le informazioni acquisite dai modelli, per definire quali canali hanno la priorità e come ordinare le campagne. Per esempio sulla base di un'analisi storica di un segmento designato, sarà più efficace mandare un'email e farla seguire da un'email diretta nelle 48 ore successive.

Risulta importante mappare i movimenti dei dati all'interno del team dei big data, affinché tutte le informazioni abbiano il loro specifico proprietario. La mappatura assicura che ogni dipendente abbia il suo determinato ruolo nel completamento del lavoro e non solo nello svolgere l'incarico individuale. E' importante inoltre che i modelli sviluppati siano applicati all'interno delle organizzazioni, per fare in modo che ciò avvenga, è molto importante dimostrare ai dipendenti più restii, il vantaggio derivante dal loro utilizzo, attraverso riunioni che esponano i successi di business che hanno adottato tali cambiamenti.

## 2.4 EFFETTI SUL SETTORE DELLA VENDITA AL DETTAGLIO

Da una ricerca svolta dal McKinsey Global Institute sul tema dei big data sono state identificate dodici leve che possono essere utilizzate lungo la catena del valore nel settore della vendita al dettaglio, queste sono state

raccolte nelle quattro principali categorie: Marketing, merchandising, operations e supply chain. In questa analisi viene stimato il potenziale impatto che ciascuna delle leve dati possono avere sul settore, si tratta per lo più di previsioni che dipenderanno anche molto da una serie di ostacoli tra cui la tecnologia, la cultura delle aziende utilizzatrici.

- Marketing

1. Cross-selling: la vendita incrociata utilizza tutti i dati che possono essere conosciuti su un cliente: la distribuzione della clientela, la cronologia degli acquisti, le preferenze, le posizioni in tempo reale e altri dati che possono incrementare gli acquisti. Amazon ad esempio utilizza filtri collaborativi per permettono di generare il suggerimento “Potrebbe interessarti anche” per ogni prodotto comprato o anche solo visitato. Questa leva è utilizzata anche per ottimizzare le promozioni che collegano prodotti complementari o in bundle.
2. Marketing posizionale: si basa sulla crescita nell'utilizzo di smartphone e altri devices che trasmettono la condivisione delle posizioni in tempo reale. Si rivolge a tutti i consumatori che si trovano nei dintorni di un negozio o all'interno di esso. Le aziende invieranno offerte speciali agli smartphone dei clienti non appena si saranno avvicinati ad i negozi. La start up PlaceCast afferma che più del 50% dei suoi clienti ha fatto un acquisto come risultato di questi annunci basati sulla localizzazione<sup>24</sup>.
3. Analisi del comportamento in-store: le analisi sul comportamento all'interno del punto vendita possono aiutare a migliorare il layout del negozio, il product mix e i posizionamenti degli scaffali. Le recenti innovazioni hanno infatti permesso ai negozianti di tenere traccia dei

---

<sup>24</sup> Abi research, Consumer Technology Barometer: Mobile,2010.



percorsi (ad esempio il tragitto e il tempo speso in uno specifico reparto) dei clienti all'interno del punto vendita.

4. Micro-segmentazione dei clienti: nonostante questa sia una leva conosciuta nel retail, i big data hanno permesso un'enorme miglioramento di questa. La quantità di dati è aumentata vertiginosamente, inoltre l'aumento nella qualità dell'analisi degli stessi ha permesso una segmentazione sempre più granulare dei clienti, al punto che alcuni rivenditori oltre ad utilizzare i dati delle ricerche di mercato e sugli acquisti storici possono anche monitorare il comportamento dei singoli clienti. Inoltre i rivenditori possono aggiornare quasi in tempo reale i dati di questi adattandoli ai cambiamenti dei clienti.
5. Sentiment analysis: tale analisi sfrutta il voluminoso flusso di dati generati dai consumatori nelle varie forme (es. nei social media) per dare supporto a una serie di decisioni di business. Ad esempio possono usare il sentiment analysis per valutare la risposta in tempo reale per le campagne di marketing e agire di conseguenza.
6. Migliorare l'esperienza multicanale del cliente: migliorare questo elemento può portare ad un incremento delle vendite, della soddisfazione del cliente e della sua fedeltà. I rivenditori possono utilizzare le informazioni raccolte (es. reddito o numero di figli) per offrire promozioni e prezzi mirati per gli acquirenti che sembrano discontinui negli acquisti.

- Merchandising

1. Ottimizzazione dell'assortimento: decidere quali prodotti portare in quali punti vendita, basandosi sui big data (es. demografie locali, percezioni degli acquirenti ecc.) può portare ad incremento reale delle vendite.
2. Ottimizzazione dei prezzi: i rivenditori oggi più che mai possono trarre vantaggio dall'incremento della granularità dei dati sui prezzi e sulle vendite insieme al più alto livello di analisi dei dati per portare l'ottimizzazione dei prezzi ad un nuovo livello. Grazie alla varietà e alla velocità dei dati è possibile comunicare e modificare i prezzi dei prodotti in tempo reale.

- Operations

1. Trasparenza delle performance: i rivenditori possono ora eseguire analisi giornalieri delle prestazioni all'interno dei punti vendita. Oggi questi sistemi riescono a fornire analisi sempre più rapide, questo permette ai managers di effettuare aggiustamenti tempestivi nelle operazioni in modo tempestivo.
2. Ottimizzazione della forza lavoro: si tratta di un'altra leva operativa che permette di creare valore riducendo i costi, mantenendo gli stessi livelli di servizio. Tale leva crea previsioni più accurate sulla necessità di lavoro, specialmente nelle ore di punta, in modo da poter ridurre l'eccesso di capacità produttiva.

- Supply chain:

1. Gestione dell'inventario: i big data aiutano a migliorare la gestione dell'inventario dei dettaglianti, attraverso sistemi di analisi dei dati i quali vengono combinati ai dati storici sulle vendite, previsioni del tempo e cicli stagionali delle vendite in

modo tale da ridurre le situazioni di esaurimento delle scorte e allo stesso modo si riduce lo spazio necessario a contenere le scorte nei magazzini, riducendo i costi di deposito delle scorte.

2. Ottimizzazione della distribuzione e della logistica: in un mondo dominato dai big data i rivenditori possono ottimizzare la fase del trasporto attraverso l'utilizzo di sistemi gps che consentono il trasferimento dei dati e l'ottimizzazione delle strade, per migliorare la produttività attraverso l'ottimizzazione del carburante, la conduzione del veicolo e la manutenzione preventiva di questo.

## 2.5 IL CASO NETFLIX

Netflix è un'azienda specializzata nella distribuzione via internet di film, serie televisive e altri contenuti di intrattenimento. Oggi l'azienda è la più grande internet tv del mondo. Essa dà la possibilità agli abbonati di poter guardare tutto ciò che vogliono in qualsiasi momento e su un'ampia gamma di dispositivi che siano dotati di connessione internet. I programmi possono essere guardati e messi in pausa, senza interruzioni pubblicitarie, in cambio di un canone mensile sottoscritto dal cliente.

## Netflix Hits 65 Million Subscribers

Number of Netflix streaming subscribers (in millions)

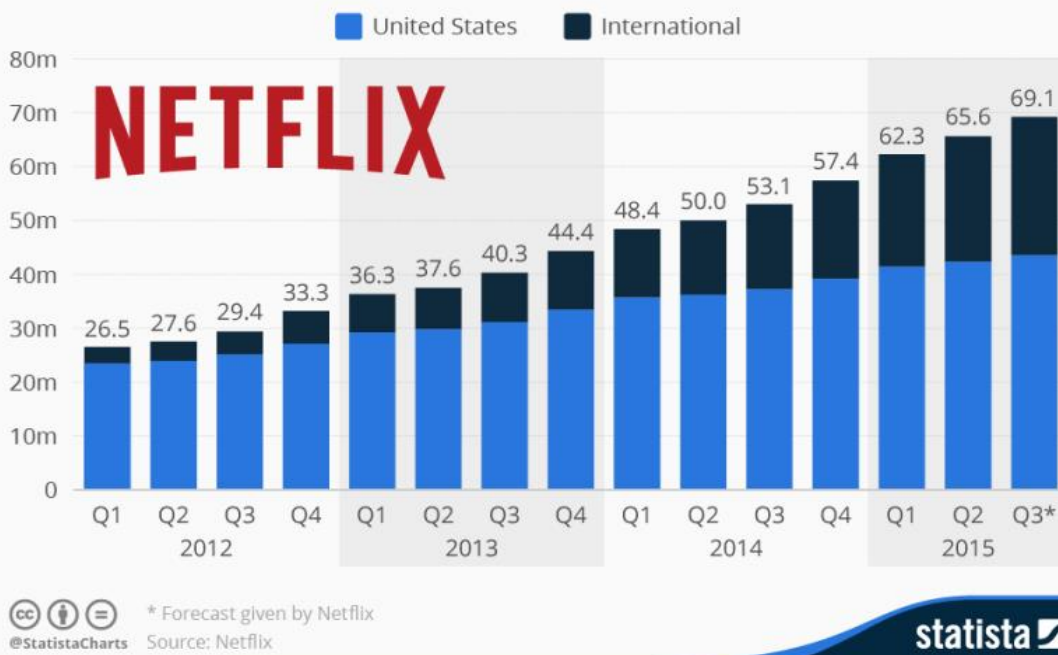


Figura 10: il numero degli utenti di netflix (2012-2015)

Risulta chiaro dall'analisi grafica che Netflix sta conquistando il mercato dell'entertainment. Dal 2012 al 2015 l'azienda ha avuto in incremento netto di 42,5 mln di utenti. Nel 2016 questi numeri sono destinati a crescere arrivando a toccare quota 45,5 mln. Questo in termini percentuali significa un incremento di circa il 100% degli utenti sottoscritti al servizio<sup>25</sup>.

I risultati economici sono di tutto rispetto e stanno confermando la l'efficacia del modello di business utilizzato dall'azienda; i ricavi risultano aumentati del 23% rispetto al dicembre dell'anno prima portando il fatturato del 2015 oltre i 7 miliardi di dollari, con un utile netto di 306

<sup>25</sup> Statista, Number of Netflix streaming subscribers worldwide from 3rd quarter 2011 to 2nd quarter 2016 (in millions)

milioni di dollari, contrazione rispetto all'anno precedente visti i continui investimenti dell'azienda (Forbes, 2016).

Come è riuscita Netflix ad ottenere questa crescita così rapida in un periodo tanto breve?

Uno dei motivi che hanno prodotto l'ottimo andamento dell'azienda è da riscontrare nel suo modello caratterizzato dall'orientamento ai big data (DDBM).

Come l'analisi dei dati ha supportato e come continua a supportare Netflix?

In primo luogo aiuta l'azienda ad ottenere informazioni rilevanti sui propri clienti, infatti i servizi analitici offrono all'azienda una serie di dati di cui necessita in modo da poter prendere decisioni migliori e più informate.

Ad oggi (2016) Netflix ha 83,18 mln di utenti in tutto il mondo, questa consistente base di utenti consente a Netflix di poter raccogliere appunto una quantità enorme di dati e rendere gli utenti più soddisfatti del servizio.

Le reti televisive tradizionali non hanno a disposizione questo tipo di privilegi, Netflix invece essendo una società sul web conosce esattamente i suoi clienti.

Ad esempio se ci si trova a guardare una delle serie televisive offerte, Netflix è in grado di quantificare quanti clienti abbiano finito di vedere la serie completa o anche quanti utenti che hanno iniziato la serie televisiva (dalla stagione 1) siano arrivati alle fine (stagione 3). Possono anche rispondere alle domande: "Qual è stato il momento in cui gli utenti hanno smesso di seguire la serie televisiva?" o "Quanto tempo è passato dal momento in cui l'utente ha terminato un episodio al momento in cui ne ha iniziato un altro?".

I dati sono raccolti in modo tale da poter capire il livello coinvolgimento dei clienti. Questi risultano essere anche più profondi rispetto a quello che risultano ad una prima analisi, Netflix traccia infatti:

- Quando si mette in pausa, si riavvolge o si avanza rapidamente
- Che giorno si guarda il contenuto (Netflix ha scoperto che tendenzialmente le persone guardano serie televisive durante la settimana e i film durante il week end.)
- La data in cui guardano i contenuti
- A che ora si guarda il contenuto
- Dove si guarda (CAP)
- Che dispositivo si utilizza per guardare
- Quando si mette in pausa e si abbandona il contenuto (e se mai tornare)
- I voti dati (circa 4 milioni al giorno)
- Le ricerche (circa 3 milioni al giorno)
- La navigazione e lo scorrimento dei programmi
- Netflix raccoglie informazione anche durante la visione dei programmi: raccoglie informazioni sul volume, i colori e gli scenari e inoltre seguono anche il cliente durante i titoli di coda.

Il motivo per cui l'azienda segue i suoi utenti anche durante lo scorrimento dei titoli di coda è perché in questo modo si possono tracciare quelli che è sono i comportamenti dopo la visione, ci si interroga su ciò che il cliente farà in seguito alla riproduzione: "lasciano l'app o tornano alla ricerca di un programma tv?" Ora Netflix sulla base delle analisi del cliente effettuate attraverso gli algoritmi specifici., manda in automatico durante i titoli di coda di un film o di una serie televisiva, la visione di un film o un episodio che siano collegati in qualche modo al precedente, in modo da mantenere alta la soddisfazione e quindi la fedeltà dell'utente.

Nel momento in cui la piattaforma acquisisce nuovi utenti chiede loro di votare qualsiasi film abbiano già visto e di valutare il loro interesse per generi di film. Questo processo, che consiste nell'aiutare i nuovi clienti a scoprire nuovi film e serie televisive, risulta essere per Netflix una delle fonti del suo successo. È importante quindi avere un algoritmo di raccomandazione, che sia piuttosto preciso in modo tale che gli utenti abbiano sempre film che vogliono vedere e in questo modo, non annulleranno il loro abbonamento al servizio.

Il sistema di raccomandazione è uno degli asset più importanti dell'azienda, risulta infatti che il 75% di quello che gli utenti guardano proviene dal sistema di raccomandazione.

Nel corso degli anni si è scoperto che gli utenti sono una fonte inestimabile di valore è importante quindi che Netflix personalizzi il più possibile il servizio di raccomandazione.



Figura 11: sistema di raccomandazione Netflix (Netflix.com)

Il sistema di raccomandazione comincia dalla home page della piattaforma, risulta infatti composta da dieci titoli che l'algoritmo del sistema suggerisce all'utente, con il termine utente Netflix identifica tutti i componenti della famiglia; ecco perché nella "top 10" dei suggerimenti risultano programmi

per i vari componenti: madre, padre e figlio<sup>26</sup>. L'algoritmo risulta in continua evoluzione, ci sono infatti intere squadre di sviluppatori che lavorano su questo, in modo tale da garantire un servizio che si adatti alle esigenze degli utenti. Vedremo ora come i big data influenzano effettivamente nelle decisioni il management della società.

Sulla base dei dati raccolti e proposti, sarebbe sbagliato affermare che l'azienda basi le sue scelte sulla base di semplici intuizioni. Sarà analizzato ora un caso pratico, di come Netflix abbia deciso di diventare produttrice di film e serie televisive, operazioni che comportano un elevato flusso in uscita e se prese in maniera sbagliata comporterebbero gravi perdite.

Nel 2011 ha preso una delle decisioni più importanti di sempre. L'azienda ha infatti scavalcato le offerte di due delle più importanti reti televisive per acquisire i diritti televisivi della versione U.S. della serie House of Cards. Spendendo circa tra i 4 e i 6 milioni di dollari per episodio, le prime due stagioni sono costate all'azienda più di 100 milioni di dollari. Uno degli investimenti più elevati mai fatti dall'azienda; ma perché un tale investimento e come l'analisi dei dati ha influenzato tale decisione?

Analizziamo il processo di analisi dei dati svolta dalla società che l'ha portata a prendere la decisione di investire sul programma.

---

<sup>26</sup> Amatriain X., Basilico J., *Netflix Recommendations: Beyond the 5 stars (Part 1)*, 2012



Guardando la cover della serie e paragonandola a quella di una serie di successo quale *Macbeth*, risultano molto simili: entrambe mostrano uomini anziani con le mani insanguinate, rispettivamente Kevin Spacey e Patrick Stewart, su sfondi neri in primo piano.

(vedi fig. 12, la ripartizione dettagliata dei colori)

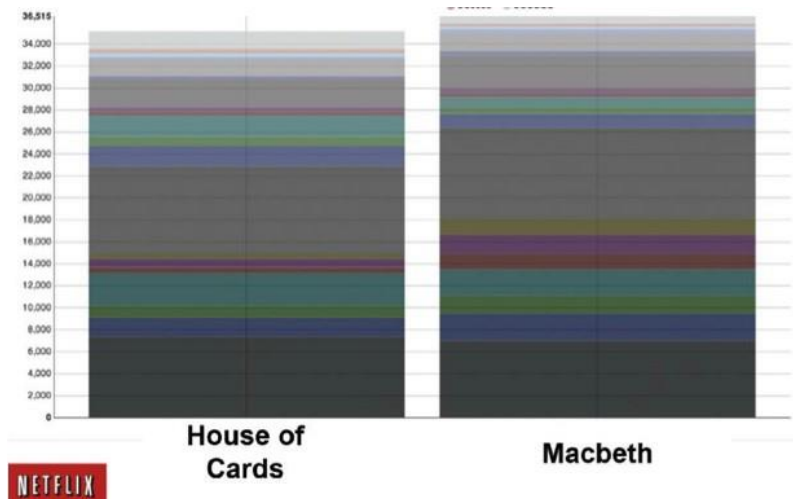


Figura 12: Comparazione dettagliata di colori: House of Cards e Macbeth

Queste analisi di colori svolte dall'azienda non sono esperimenti occasionali, ma sono il risultato di lunghe ricerche. Netflix riconosce infatti che esiste un enorme potenziale all'interno di queste scoperte. Quindi le scelte non risultano mai casuali, ma avvengono sempre in risposta a “domande” poste ai dati raccolti come per esempio “Quali sono i colori che attirano di più lo sguardo del cliente?” ecc.<sup>27</sup>

Prima che venisse dato il via libera per l'acquisizione del programma, l'azienda aveva una serie di dati raccolti tramite la piattaforma, che potevano essere utili: molti degli utenti hanno guardato il film “The Social Network” diretto da David Fincher dall'inizio alla fine. La serie inglese di “House of Cards” ha avuto un buon seguito ed inoltre gli stessi che hanno guardato la serie hanno anche guardato film che avevano come protagonista Kevin

<sup>27</sup> Simon P., Wired: *Big Data Lessons from Netflix*, 2014.

Spacey diretti dal regista David Fincher. Ovviamente questi tre fattori sinergici contenevano un elevato volume di utenti al loro interno.

La combinazione di questi fattori, ha avuto un grande peso nella decisione di Netflix di investire nella creazione della versione americana di “House of Cards”; dalla dichiarazione di Swasey S., VP delle comunicazioni sociali, si ricava che, l’azienda attraverso lo studio dei dati e dalla combinazione di questi, ha potuto prendere la propria decisione di investimento facilmente.

Dopo aver effettuato l’investimento Netflix ha svolto una campagna promozionale caratterizzata da “trailer personalizzati” per ogni tipo di utente della piattaforma. L’azienda ha infatti prodotti dieci diversi tagli di trailer della serie televisiva, ciascuno orientato ad un tipo di pubblico diverso, il trailer che ogni utente vedeva era scelto sulla base delle sue visualizzazioni. Ad esempio se l’utente aveva visto molti film con Kevin Spacey il trailer sarebbe stato caratterizzato in buona parte dalla presenza dell’attore, allo stesso modo se l’utente aveva visto molti spettacoli con protagoniste donne, avrebbe visto un trailer con le donne protagoniste della serie televisiva.

Sulla base di queste analisi si può concludere che la serie proposta dall’azienda sia stata un successo?

Tipicamente il tasso medio di successo di un programma televisivo dopo che è stato lanciato in rete è del 35% in media<sup>28</sup>.

Nel 2013 Netflix disponeva di sette spettacoli televisivi, cinque dei quali erano stati rinnovati per altre stagioni, il tasso di successo dell’azienda si attesta quindi su un 70%. Si può quindi affermare legittimamente che i dati abbiano fatto la differenza.

House of Cards ad esempio ha portato oltre due milioni di nuovi iscritti solo negli Stati Uniti, alla piattaforma nel primo trimestre del 2013, insieme ad

---

<sup>28</sup> Ocasio A., Screenrant.com: *Tv Success Rate: 65% of New Shows Will Be Canceled*, (2012)

un altro milione di nuovi abbonati da altre parti del globo. Come riportato dal magazine statunitense Atlantic Wire, questi 3 milioni di abbonati hanno quasi ripagato alla società il costo della produzione della serie.

La serie oltre a portare nuovi clienti ha anche fatto in modo che i clienti già iscritti al servizio rimanessero fedeli, è stato dimostrato infatti che l'86% dei clienti è rimasta fedele a causa di House of Cards. Si può intuire, a questo punto che Netflix basa le sue scelte su un'attenta analisi dei dati.

L'azienda non propone esclusivamente film autoprodotti, bensì la maggior parte delle proposte che si trovano sono film acquisiti in licenza, e acquisirle risulta essere un'attività molto costosa. È necessario per il management trovare quei film che possano garantire il più alto livello di audience rispetto ai costi sostenuti per acquistare le licenze degli stessi; sono utilizzati gli stessi metodi di analisi dei dati attraverso cui si cerca di carpire ciò che gli utenti amano guardare rispetto a ciò che è disponibile su Netflix. (McCabe J., Director of Global Media Relations).

La società deve ponderare le proprie scelte rispetto a ciò che gli utenti guardano e ciò che vorrebbero vedere, essa infatti non può semplicemente acquistare ogni colossal che viene lanciato al cinema, ma deve essere in grado di acquisire il massimo vantaggio dall'analisi dei dati e rendere gli utenti soddisfatti mantenendo un livello di costi che sia efficiente. Ad esempio il film "Batman: Il Cavaliere Oscuro" è stato uno dei film più popolari del 2008, con un risultato di oltre un miliardo di dollari al botteghino. Netflix sapeva che il film avrebbe soddisfatto i suoi utenti, ma lo studio produttore del film richiedeva un prezzo molto elevato per l'acquisizione della licenza<sup>29</sup>. La società ha ponderato la possibilità di effettuare l'acquisto di tale titolo con la possibilità di acquistare invece i diritti di altri sei film con gli stessi attori e

---

<sup>29</sup> Bulygo Z., Kissmetrics.com: *How Netflix Uses Analytics To Select Movies, Create Content*. (2013)

regista, chiedendosi quale scelta avrebbe portato alla maggiore soddisfazione degli utenti per la stessa fuori uscita di denaro.

Dall'analisi si può vedere come Netflix prenda decisioni consapevoli sulla base dei dati.

Ovviamente i dati da soli non possono portare sempre a soluzioni ottimali, ci sono alcune situazioni in cui l'intuizione deve prendere il sopravvento. Quello che i dati e la loro analisi può fare però è dare una visione delle cose in modo da poter rendere migliore il business e offrire un prodotto di maggior qualità. Le società che possiedono i dati ed hanno la capacità di analizzarli possiedono un vantaggio rispetto a quelle che basano le loro scelte esclusivamente sulle intuizioni.

## CONCLUSIONE

Le aziende, attraverso la giusta applicazione dei big data, avranno la possibilità di migliorare l'efficienza operativa, ridurre i costi, migliorare le performance produttive, le relazioni con i clienti, accelerare le consegne, formulare e rispondere a richieste più approfondite, migliorare e semplificare il processo decisionale. Le aziende si troveranno quindi di fronte ad una vera e propria innovazione. Questa attività richiederà però del tempo, elevati investimenti in tecnologie e know-how altamente specializzato. L'impatto finale però sarà un vantaggio competitivo, nel lungo termine, che difficilmente potrà essere recuperato dai competitors che non sapranno trarre i giusti benefici dai dati.

Un numero sempre più considerevole di aziende sta seguendo le orme di Netflix e delle altre imprese "pioniere", intravedendone gli innumerevoli vantaggi. Netflix sta incrementando in modo sorprendente il numero dei suoi iscritti, sfruttando le analisi ricavate dallo studio dell'ingente mole di dati a sua disposizione. Sicuramente sarà più agevole per le nuove imprese, strutturare la loro organizzazione in modo da riuscire a sfruttare al meglio i dati a disposizione e quindi massimizzare i loro profitti. Le imprese che sono già attive sul mercato saranno costrette a cambiare la loro struttura organizzativa. Questo porterà conseguentemente ad un aumento dei costi e al sorgere di difficoltà interne, ad esempio dovranno essere affrontati elevati costi per la formazione del personale e per il miglioramento della tecnologia in modo da non rimanere indietro rispetto alla concorrenza, sempre maggiore in questo campo.

All'interno dell'elaborato, analizzando un notevole numero di fonti e dati, ho cercato di fornire al lettore una visione generale sul mondo dei big data e, procedendo nella stesura del testo, di captare i vantaggi che tale innovazione potrà apportare. Nella prima parte ho cercato di portare alla luce le cause che

hanno comportato la nascita dei big data. I dati oggi risultano molto più numerosi, ma non solo. È possibile infatti elaborarli e conservarli ad un costo che risulta essere irrisorio rispetto al passato. Questo fenomeno porterà ad un cambiamento epocale, sarà infatti possibile espandere il campione in analisi e non limitarlo ad un campo ristretto come veniva fatto in precedenza, garantendo analisi più precise e in più ambiti. Il fenomeno è ancora in fase di decollo, non esiste ancora una definizione che possa essere chiara e precisa di quello che i big data possano rappresentare. Come riportato in conclusione del primo capitolo, i big data sono stati paragonati a nuove tecnologie, che possono essere ben rappresentate dalla curva ad “S” del cambiamento tecnologico. In questo modo si è messo in risalto il fatto che ad oggi i big data presentano un margine di espansione ancora maggiore, le performance garantite dalla tecnologia saranno superiori. Tale fenomeno però comporterà anche delle problematiche: in primo luogo, come dimostrato dal caso Google “G.F.T.”, un’analisi che si basa solo sui dati escludendo ogni fondamento teorico, risulta essere fragile. È necessario che i dati siano sostegno di un’analisi che sia in primo luogo basata su studi e conoscenze del settore considerato. Un’altra problematica analizzata dall’elaborato è la privacy, che risulta essere inversamente correlata all’espansione dei dati. Sarà necessario quindi che i governi e le istituzioni garantiscano ai cittadini e potenziali clienti, di mantenere inalterato il loro status e fare in modo che non vi siano sfruttamenti illeciti da parte delle organizzazioni più forti sul mercato. Una soluzione interessante potrebbe essere quella di garantire al cittadino il possesso delle proprie informazioni, al tempo stesso dando la possibilità a chi lo desidera di poter immettere sul mercato le proprie informazioni, generando così un libero mercato dei dati dei singoli cittadini.

Nella seconda parte della tesi sono stati raccolti dati, con lo scopo di dimostrare che il mercato dei big data risulta essere nuovo e soprattutto profittevole. Dall’esposizione si evincono quelli che sono i settori che utilizzano un maggior numero di dati, tra questi si ritrova il settore della

vendita al dettaglio. Nel paragrafo 2.4 sono stati analizzati gli effetti che i dati porteranno in tale settore e le possibili barriere che si verranno a generare.

Le aziende avranno a disposizione un'ingente mole di dati, ma una delle maggiori difficoltà sarà appunto quella di riuscire a capitalizzarle, in particolare adeguando la loro struttura al fine di poter supportare l'orientamento ai dati. Nel paragrafo 2.2 è infatti analizzato il modello *Data-driven Business Model*, che presenta delle peculiarità utili a garantire il miglior utilizzo e applicazione dei dati generati e raccolti dalle aziende. In ultima analisi è stato approfondito il caso Netflix, società leader dell'internet tv. Si è voluto infatti dimostrare attraverso casi pratici come, gli effetti positivi scaturiscano da un utilizzo proficuo dei dati insieme a strutture organizzative ad hoc. La compresenza di questi tre fattori in sinergia tra loro, ha garantito all'azienda ottimi risultati e le previsioni risultano ancor più vincenti.

## BIBLIOGRAFIA/SITOGRAFIA

- Hilbert, M., How much information is there in the “information society”? *Significance*, 9(4), 8–12.
- Prospetto IPO di Facebook, <<Form S-1 Registration Statement>>, U.S. Securities and Exchange Commission, 1° febbraio 2012.
- T. Geron, Twitter’s Dick Costolo: Twitter Mobile Ad Revenue Beats Desktop on Some Days, <<Forbes>>, 6 giugno 2012.
- Mayer Schönberger V., Cuckier K, *Big Data: una rivoluzione che trasformerà il nostro modo di vivere e già minaccia la nostra libertà*. Milano: Garzanti, 2013
- Ibrahim Abaker Targio Hashem, *Information Systems journal: The rise of “big data” on cloud computing: Review and open research issues* (2014).
- Mayer Schönberger V., Cuckier K, *Big Data: una rivoluzione che trasformerà il nostro modo di vivere e già minaccia la nostra libertà*. Milano: Garzanti, 2013
- <https://techcrunch.com/2008/04/17/microsoft-acquires-forecast-for-115m/>
- <sup>1</sup>Giuseppe Roccasalva, 201, p. 97, “I Big Data e gli strumenti di visualizzazione analitica: interazioni e studi induttivi per le P.A.”.
- Fontana F., Caroli M.: *Economia e Gestione delle Imprese*, McGraw-Hill, (2013). Cap 11.
- A.F. Dugas et al., Google Flu Trends: Correlation with Emergency Department Influenza Rates and Crowding Metrics; *CID Advanced Access*, 8 gennaio 2012
- Harford T., *Financial Times*, “Big data are we making a big mistake?”, 28 marzo 2014.
- ARRINGTON M., *TECHCRUNCH: FIRST PERSON IDENTIFIED FROM AOL DATA: THELMA ARNOLD*, 2006.
- DAVIS J., *INFORMATIONWEEK: BIG DATA, ANALYTICS SALES WILL REACH \$187 BILLION BY 2019*, MAGGIO 2016
- Press G., *Forbes: 6 Predictions For Big Data Analytics And Cognitive Computing In 201*, ottobre 2015
- Hartmann P., Zaki M., Fieldmann N., Neely A., University of Cambridge: *Big Business? A taxonomy of Data-driven Business Models used by Start-up firms*, marzo 2014.
- Rayport J., Sviokla J., *HBR: Exploiting the virtual chain value*, novembre-dicembre 1995.
- E. Brynjolfsson; L. Mitt; H. Kim, *Social science research network paper: “Strength in Numbers: How does Data-driven Decision Making Affect Firm Performance?”* (2011)
- Aricker M., McGuire T., Perry J., *Harvard Business Review: Five Roles You Need on Your Big Data Team*, luglio 2013.
- Manyika J., Chui M., Brown B., *Mckinsey Global Institute: Big Data: The next frontier for innovation, competition, and productivity*, maggio 2011.



Abi research, Consumer Technology Barometer: Mobile,2010.

Statista, Number of Netflix streaming subscribers worldwide from 3rd quarter 2011 to 2nd quarter 2016 (in millions)

Amatriain X., Basilico J., Netflix Recommendations: Beyond the 5 stars (Part 1), 2012

Simon P., Wired: Big Data Lessons from Netflix, 2014.

Ocasio A., Screenrant.com: *Tv Success Rate: 65% of New Shows Will Be Canceled*, (2012)

Bulygo Z., Kissmetrics.com: *How Netflix Uses Analytics To Select Movies, Create Content*. (2013)