



UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Scienze Economiche e Aziendali “M. Fanno”

Corso di laurea triennale in Economia
Prova Finale

BIG DATA VETTORE DI CAMBIAMENTO

IL CASO NETFLIX

Relatore
Prof. Luisa Bisaglia

Laureanda
Allegra Belluco
Matricola: 1141560

Anno accademico 2017/2018



Fonte: Google

“Big Data is like a teenage sex: everyone talks about it, nobody really knows how to do it, everyone thinks everyone else is doing it , so everyone claims they are doing it...”

(Dan Ariely, Duke University)

Sommario

CAPITOLO 1. Il mondo dei <i>Big Data</i>	7
1.1 L'individuo come fonti di dati	7
1.2 Le caratteristiche dei big data	8
1.2.1 <i>I soggetti attivi</i>	8
1.2.2 <i>Le 4V</i>	9
1.3 La catena del valore (il ciclo di vita del dato)	12
1.4 L'industria audiovisiva e il cambiamento rispetto al modello tradizionale	14
CAPITOLO 2. Analisi del caso Netflix e della sua strategia di business	16
2.1 La storia di Netflix: il suo debutto nel mercato televisivo online	16
2.1.1 <i>Netflix abbandona il feedback a stelline e passa ai giudizi con il pollice di Facebook</i>	19
2.2 In che modo Netflix raccoglie e analizza i dati dagli utenti. E che dati raccoglie?	20
2.3 Big Data per l'analisi dei contenuti	22
2.3.1 <i>Esiste un team di lavoro specializzato</i>	22
2.3.2 <i>'House of Cards': prima serie originale Netflix di successo</i>	23
2.4 Introduzione ai differenti algoritmi	26
2.4.1 <i>'Personalized Video Ranker'</i>	27
2.4.2 <i>'Top-N'</i>	27
2.4.3 <i>'Trending Now'</i>	28
2.4.4 <i>'Because You Watched'</i>	28
2.4.5 <i>L'algoritmo di generazione della pagina</i>	28
2.4.6 <i>Vecchi e nuovi membri</i>	30
2.5 A livello internazionale	31
2.6 Una prospettiva leader-follower	34
CAPITOLO 3. I rischi legati alla raccolta e all'utilizzo dei Big Data	36
3.1 Violazione della Privacy	37
3.1.1 <i>Asimmetrie informative</i>	38
3.1.2 <i>Misure per la protezione dagli abusi</i>	39
3.2 Altri rischi relativi allo scambio di dati	39
3.1.1 <i>Discriminazioni</i>	40
3.1.2 <i>Disinformazione e false interpretazioni dei dati</i>	40
3.1.3 <i>Sul caso Netflix</i>	40
Conclusioni	42
➤ Analisi SWOT:	42
➤ Cosa imparare da Netflix	43
Bibliografia	44

Introduzione:

Con il presente elaborato cercheremo di esplorare il mondo dei *big data*, fenomeno che sta modificando il focus all'interno delle aziende. In particolare, entreremo nello specifico delle imprese televisive, le quali rappresentano uno dei casi di maggior successo riguardo all'analisi e all'utilizzo dei *big data*.

Oggi ci troviamo in un'epoca in cui diventa sempre più fondamentale l'utilizzo dei dati nei diversi processi decisionali sia di imprese sia di noi come cittadini. Esso ha dato il via a un radicale cambiamento che coinvolge sia aspetti economici sia l'intera società. Le tecnologie che abbiamo a disposizione permettono di convertire in formato digitale - cioè in dati - ogni cosa, dalla più banale alla più inimmaginabile, come ad esempio i pensieri o i movimenti del corpo.

Tutto viene quindi trasformato in dati per mezzo di fonti che possono essere qualsiasi tipo di device, sensore, social network, o sistema operativo.

“Big Data è un termine che si riferisce ad un ampio volume di dati, strutturati e non, che sommergono quotidianamente un'azienda. Ciò che conta però non è la quantità di dati, ma come vengono utilizzati: possedere big data significa analizzarli per ottenere le informazioni necessarie a prendere le migliori decisioni aziendali.”

(SAS, 2018. *Big data. Cosa sono e perché sono importanti*)

Il fenomeno che rappresenta la trasformazione di qualsiasi informazione in dato è detto "datizzazione" (Agcom, 2018, p.2) ed è proprio il progressivo aumento dell'uso di strumenti moderni e innovativi di comunicazione elettronica e la conseguente digitalizzazione dei processi produttivi che danno origine ad un ammontare immenso di dati economici e sociali, disponibili ed elaborati ad una velocità sempre maggiore ma anche a una crescente varietà di formati.

La principale inesauribile sorgente di dati è costituita dai cosiddetti dispositivi mobili, che ognuno di noi utilizza ogni giorno in maniera più o meno frequente. Quelle che noi lasciamo mediante questi dispositivi, sono dette *online footprint*, ovvero tracce che rimangono memorizzate nel sistema sia quando noi ci spostiamo da un luogo ad un altro, sia quando effettuiamo pagamenti online, sia quando condividiamo il nostro pensiero mediante un commento o postiamo una foto.

Social network e motori di ricerca, come molte altre piattaforme, raccolgono e utilizzano tutti questi dati come asset strategico per la loro nuova offerta di servizi e contenuti e per migliorare quelli esistenti, cercando anche di innovare i processi produttivi e distributivi per andare incontro alle esigenze di consumatori e cittadini.

“I grandi collettori e utilizzatori di big data sono le società che offrono servizi via web a grandi masse di utenti come Google, Apple, Facebook, Amazon, Netflix, ma un numero crescente di imprese nel campo dei media come, più in generale, in molte altre attività economiche, utilizzano nelle proprie strategie produttive e commerciali il trattamento di big data.” (Richeri, 2016, p.2)

L'utilizzo senza dubbio innovativo che oggi viene fatto dei dati da parte degli esseri umani nei loro processi decisionali costituisce un radicale cambiamento rispetto a quanto avveniva nel passato e coinvolge tutte le società che abbiano apportato al loro interno le nuove tecnologie e le nuove tecniche di archiviazione per la raccolta e l'analisi di questi dati. Tutto questo sta avendo un impatto diretto su quelle che sono le abitudini di noi individui e quindi anche sulla qualità della nostra vita e sulle opportunità per il nostro futuro.

In particolare, nel mio elaborato, ho deciso di approfondire la parte che riguarda il mondo della televisione poiché le imprese che in questo campo fanno utilizzo dei big data hanno riscontrato dei risultati molto positivi e hanno avuto un successo davvero notevole. Il merito è tutto del nuovo modello commerciale che queste imprese utilizzano, che è completamente diverso da quello tradizionale.

Le imprese coinvolte, infatti, non offrono più contenuti televisivi prefissati in una programmazione standardizzata e identica per tutti gli utenti, ma offrono, invece, contenuti accessibili a richiesta (video on demand) e selezionabili a seconda delle diverse preferenze. Questo rappresenta un cambiamento radicale nel mercato televisivo mondiale in quanto il cliente ora è portato a scegliere autonomamente il programma desiderato, nel momento che più preferisce e tra un elenco di titoli molto vasto, pagando questo servizio secondo diverse modalità.

Ovviamente, alla base di tutto questo cambiamento c'è l'utilizzo di Internet, senza il quale non sarebbe stato possibile raggiungere livelli tanto elevati di successo per l'offerta televisiva.

I vantaggi però non vanno solo in tasca al cliente del servizio, ma anche delle imprese in gioco, le quali, oltre a offrire un servizio innovativo e molto attraente, si arricchiscono di informazioni sui propri clienti che possono riguardare sia i loro gusti sia tutto ciò che riguarda il consumo che questi ne fanno. Si abbandonano quindi le vecchie stime che si facevano col modello lineare e standardizzato di prima (Richeri, 2016) e si giunge a una conoscenza precisa e puntuale del

comportamento di ciascun telespettatore che fa uso di questi nuovi modelli televisivi. Nel CAPITOLO 2 verrà analizzato il caso di Netflix, che costituisce un chiaro esempio di successo nel campo televisivo dei big data.

Oltre agli aspetti positivi e all'elevato grado di innovazione e potenzialità che questo nuovo approccio apporta all'intero sistema, sorge tuttavia, come vedremo, la necessità di far fronte al trade-off che si presenta 'tra il valore commerciale dell'informazione e il rispetto di diritti individuali e collettivi fondamentali quali la privacy, la tutela della concorrenza e le garanzie del pluralismo informativo' (Agcom, 2018). Vi sono infatti dei rischi a proposito che vanno sottolineati e costituiscono il rovescio della medaglia del notevole successo dei big data (come vedremo nel CAPITOLO 3). Alcuni dei rischi che verranno analizzati sono, per esempio, la violazione della privacy, la disinformazione online, le asimmetrie informative tra consumatori e gli operatori dei servizi online, l'incompletezza dei contratti che disciplinano i diritti di proprietà sui dati e l'assenza di espliciti mercati che regolino la formazione dei prezzi. Tutte queste sono cause che portano a veri e propri fallimenti di mercato.

CAPITOLO 1. Il mondo dei *Big Data*

1.1 L'individuo come fonti di dati

Al giorno d'oggi, l'impiego dei *big data* è diventato senza dubbio pervasivo nella vita quotidiana di ognuno di noi. Esso si estrae da un numero sempre più crescente di settori: dall'energia alla finanza, dal commercio elettronico all'intrattenimento, dalle assicurazioni all'*automotive* e così via.

Enormi masse di dati, infatti, con strutture tra di loro differenti, vengono raccolte e utilizzate per diverse finalità mediante l'utilizzo di tecnologie e processi sempre più sofisticati.

Tale fenomeno riguarda, sia a livello nazionale che a livello internazionale, tutte le piattaforme online, che si stanno diffondendo a ritmo incalzante tra la popolazione. Stiamo parlando di social network, ma anche di motori di ricerca e operatori online di vario genere che assumono un ruolo molto rilevante nella vita di ogni individuo. Ognuno di noi, giornalmente, compie azioni mediante un qualche strumento che abbia connessione a Internet. Ognuno di noi, quando naviga in rete, comunica qualcosa di sé e delle sue preferenze, anche inconsciamente.

La capacità di accedere ai dati che riguardano i propri utenti e di utilizzarli come asset strategico per il raggiungimento dei propri obiettivi aziendali (come la personalizzazione dell'esperienza per l'utente in rete e la massimizzazione dei profitti), rappresenta un elemento fondamentale nel modello di business delle piattaforme online. Ed è proprio per questo che è un argomento che merita di essere approfondito e discusso.

L'ecosistema informativo mondiale sta cambiando progressivamente la sua struttura in seguito alla diffusione sempre più incontrollabile di tutte le informazioni che girano in rete, per le molteplici azioni che gli utenti compiono navigando online, e per le emozioni e i pensieri che esprimono attraverso le proprie pagine social. La ricerca o la visualizzazione di contenuti in Internet rende l'individuo un generatore di dati. Sono proprio i social network e i motori di ricerca, infatti, che in cambio di offrire un servizio senza un corrispettivo di prezzo per l'utente finale, recuperano e analizzano tutta questa enorme quantità e varietà di dati sugli individui.

Per fare ciò, le piattaforme online, si servono di appositi algoritmi, ovvero “sottoinsiemi logici fondati su funzioni matematiche presenti in varie tipologie di software” (Agcom, 2018, p.15). Questi algoritmi sono pertanto determinanti nel definire le modalità di consumo informativo

degli utenti, e assumono un valore significativo anche dal lato dell'offerta nel proporre le migliori soluzioni all'utente e per determinare le scelte di business delle diverse aziende.

1.2 Le caratteristiche dei big data

1.2.1 I soggetti attivi

“Nell'ecosistema dei *big data*, è possibile identificare i seguenti attori principali:

- i **soggetti generatori di dati** (o fornitori di dati);
- i **fornitori della strumentazione tecnologica**, tipicamente sotto forma di piattaforme per la gestione dei dati;
- gli **utenti**, cioè coloro che utilizzano i *big data* per creare valore aggiunto. Infatti, “lo sviluppo vertiginoso della strumentazione tecnologica per la raccolta, conservazione, classificazione e processamento dei dati”, ha generato la necessità di avvalersi di personale specializzato, in grado di estrarre valore dai dati, e di capacità legate alla raccolta e conservazione di questa ingente mole di informazioni. È importante sottolineare quindi che “le professionalità che consentono di estrarre valore dai big data, rappresentano un'importantissima leva competitiva per le imprese, che hanno bisogno della disponibilità di servizi di storage e scalabilità dei dati (Agcom 2018).”
- i **data brokers**, cioè le organizzazioni che raccolgono dati da una varietà di fonti sia pubbliche, sia private, e li offrono, a pagamento, ad altre organizzazioni;
- le **imprese e le organizzazioni di ricerca**, la cui attività diventa fondamentale per lo sviluppo di nuove tecnologie, di nuovi algoritmi attraverso cui esplorare i dati ed estrarre valore;
- gli **enti pubblici**, sia in qualità di enti regolatori dei mercati, sia con riferimento alle attività della pubblica amministrazione volte a migliorare i prodotti e i servizi offerti alla cittadinanza e in grado di aumentare il benessere collettivo.” (Agcom, 2018, p.3)

Le interrelazioni tra i diversi soggetti aumentano la complessità dell'intero sistema e fanno sì che all'interno di questo scenario operino (poche) grandi imprese multinazionali “caratterizzate da un elevato grado di integrazione in tutte le fasi dell'ecosistema, accanto a una miriade di piccole imprese specializzate.” (Agcom, 2018)

1.2.2 Le 4V

Il maggior punto di forza del modello risiede nelle 4V, che sono: volume, varietà, velocità e valore. Queste variabili, individuate da Gartner Inc. (società per azioni leader mondiale nella consulenza strategica, ricerca e analisi nel campo dell'Information Technology)¹, aumentando di giorno in giorno con un'intensità senza precedenti, risulterebbero molto difficili da analizzare con i metodi tradizionali di analisi dei dati, obsoleti e inadeguati.

A questo proposito, quindi, vengono utilizzate apposite tecniche di analisi dei dati che permettono la gestione di questo nuovo complesso sistema, che risulta essere molto differente da quello tradizionale.

Analizziamo più in dettaglio le 4V:

- 1) **Volume.** Rappresenta la caratteristica che più intuitivamente si può associare ai big data visto il riferimento implicito alla dimensione del fenomeno. Si tratta infatti di una dinamica di crescita di dati esponenziale.

“La dimensione si misura in terabytes (mille gibabytes), petabytes (un milione di gigabytes), zettabytes (un milione di petabytes) che solo grandi apparati digitali sono in grado di memorizzare ed elaborare, mentre i 'little data' sono quelli che si misurano in gigabyte (mille megabyte) che un normale personal computer è in grado di memorizzare e trattare, così come un tablet o uno smartphone. Per esempio registrare 7 minuti di televisione in alta definizione richiede un gigabyte di memoria mentre con un petabyte si possono registrare 13,3 anni di trasmissioni televisive in alta definizione.” (Richeri, 2016)

Quindi fondamentale, per l'analisi di questi dati, è la rapidità con cui viene aggiornata l'unità di misura dell'informazione.

¹ Ibrahim Abaker Targio Hashem, 2014.: “The rise of “big data” on cloud computing: Review and open research issues”. Information Systems journal

Unità e simbolo	Dimensione	In termini pratici ...
Bit (b)	1 o 0	L'unità di misura elementare dell'informazione, diminutivo di " <i>binary digit</i> " che viene rappresentata alternativamente con le cifre 0 e 1, in quanto corrisponde a una scelta tra due alternative egualmente possibili.
Byte (B)	8 bits	L'informazione utile a creare un singolo carattere nel codice del computer; è l'unità di base di calcolo.
Kilobyte (KB)	1.000 o 210 bytes	Una pagina di testo equivale a 2KB. 100KB misurano una immagine fotografica in bassa risoluzione.
Megabyte (MB)	1.000KB o 220 bytes	Un file MP3 di un brano musicale "tipo" è pari a 4MB. 100MB equivalgono ad una pila di libri pari ad un metro.
Gigabytes (GB)	1.000MB o 230 bytes	Un film della durata di circa due ore può essere compresso in 1-2GB. Un testo di 1GB contiene all'incirca 1 miliardo di caratteri, ovvero circa 4.500 libri di 200 pagine o 240.000 caratteri.
Terabyte (TB)	1.000GB o 240 bytes	1TB equivale a 262.144 file MP3 (con una durata media di 4 MB). 1TB equivale a circa 4.580.000 libri di 200 pagine. Tutti i libri catalogati nella <i>America Library of Congress</i> ammontano a 15TB. Tutti i <i>tweet</i> inviati prima del 2013 equivalgono ad un file di testo di 18.5TB; per stampare un testo simile (ad una velocità di 15 pagine formato A4 per minuto) ci vorrebbero 1200 anni.
Petabyte (PB)	1.000TB o 250 bytes	1PB corrisponde a circa 4.691.000.000 libri da 200 pagine. La NSA (<i>National Security Agency</i>) analizza circa l'1,6% del traffico internet globale, circa 30PB al giorno. Volendo ascoltare 30PB di musica senza soluzione di continuità, ci vorrebbero più di 60.000 anni.
Exabyte (EB)	1.000PB o 260 bytes	1EB di dati corrisponde ad una capacità di immagazzinamento dei dati corrispondente a 33.554.432 di <i>iPhone 5</i> con una memoria di 32GB. Nel 2018, il volume mensile del traffico dati tramite telefonia mobile si stima ammonti a 1EB; se questo ammontare di dati fosse conservato in smartphone <i>iPhone 5</i> da 32 GB, sarebbe necessario formare una pila di <i>iPhone 5</i> alta 239 l'Empire State Building.
Zettabyte (ZB)	1.000EB o 270 bytes	1ZB corrisponde a 281.474.977.500.000 file MP3 della grandezza media di 4MB, ovvero 250.000.000.000 di DVD da 4,38 GB.
Yottabyte (YB)	1.000ZB o 280 bytes	Il contenuto di codice genetico appartenente ad una singola persona può essere immagazzinata in meno di 1,5GB; ciò implica che 1YB può contenere il genoma di 800 trilioni di individui, ovvero 100.000 volte circa la popolazione del pianeta.

FIGURA 1. Le unità di misura dell'informazione

(Fonte: Elaborazione dell'Autorità su dati Economist, www.computerhope.com, Cisco e Emmanuel Letouzé)

L'enorme quantità di dati in circolazione implica, a volte, anche la ridondanza di essi, ma in questo caso 'ridondanza' non è sinonimo di 'inutilità'. Basti pensare che grazie a questi dati, società di servizi web come per esempio Google, sono state in grado di perfezionare sempre di più il proprio motore di ricerca, anche se a prima vista le informazioni prese singolarmente, sembravano assolutamente prive di importanza. È importante sottolineare, però, il fatto che al crescere del volume dei dati, aumentano i costi per l'archiviazione e conservazione di questi. Non solo, aumentano anche quelli legati all'estrazione di valore del dato, "dal momento che si rende necessario l'utilizzo di sofisticati algoritmi e software e di figure professionali altamente specializzate e in grado di gestire la complessità." (Agcom, 2018)

- 2) **Varietà.** L'eterogeneità dei dati sta crescendo in maniera esponenziale di giorno in giorno. La varietà si riferisce sia al tipo di fonte da cui viene recuperato il dato, sia al formato con cui vengono acquisiti e sotto che forma essi vengono rappresentati. La gran

parte di questi dati vengono chiamati ‘non strutturati’ in quanto non sono organizzati secondo una precisa struttura e quindi hanno bisogno di essere elaborati mediante complesse tecniche di elaborazione per trarne risultati che possono poi essere utilizzati per i diversi fini aziendali. È evidente quindi la complessità nella gestione di questa eterogeneità.

- 3) **Velocità.** Essa è connessa alla altissima frequenza con cui i dati giungono dalla loro sorgente, al punto dove vengono raccolti. Questo avviene grazie anche, e soprattutto, alle innovazioni tecnologiche per la raccolta di questi dati sempre in continua crescita. È importante che sia presente la velocità nel processo di raccolta e elaborazione poiché solo in questa maniera le imprese riescono a rispondere prontamente alle richieste del pubblico e ai cambiamenti che riguardano le sue preferenze e necessità.

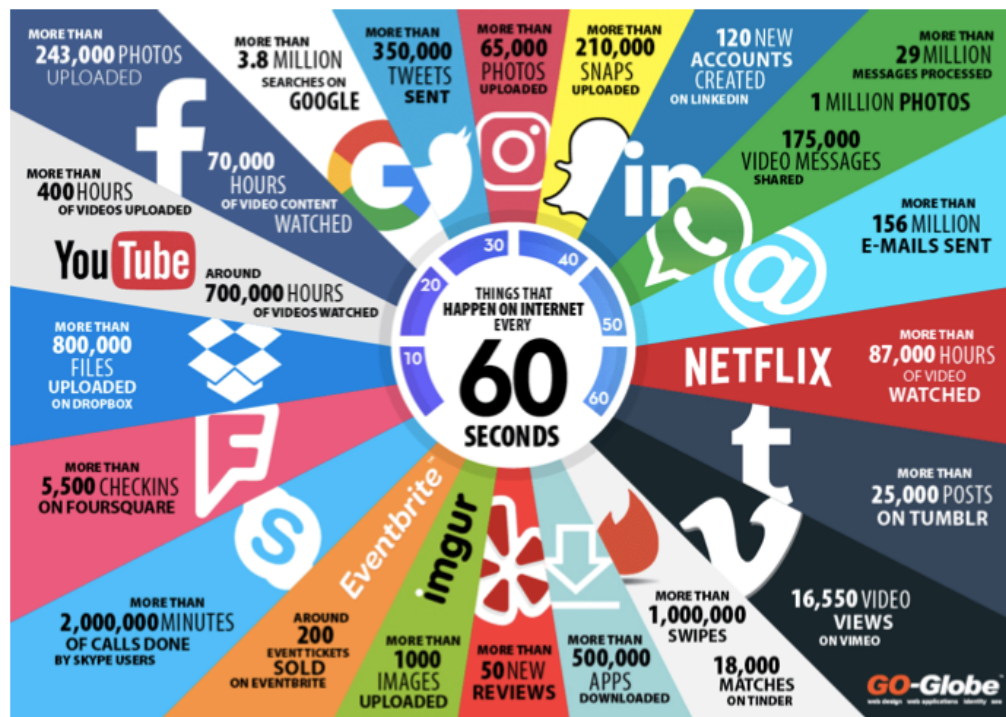


FIGURA 2. Big data: Volume, Varietà e Velocità dei dati (il flusso di dati su internet in 60 secondi)

- 4) **Valore.** 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. Essi non sono necessari solo per generale valore per le imprese, ma anche per accrescere il bisogno complessivo della società. È fondamentale quindi valutare la veridicità e la qualità dei dati poiché solo lavorando attraverso dati attendibili le analisi possono portare a risultati pertinenti (Signore, 2016).

1.3 La catena del valore (il ciclo di vita del dato)

I dati, una volta che vengono raccolti mediante gli appositi mezzi, entrano in una serie di fasi che consentono che questi vengano analizzati e dotati di valore per rispondere ai fini dell'azienda che ha messo in atto il processo. Infatti, presi singolarmente, questi dati non danno nessuna informazione rilevante.

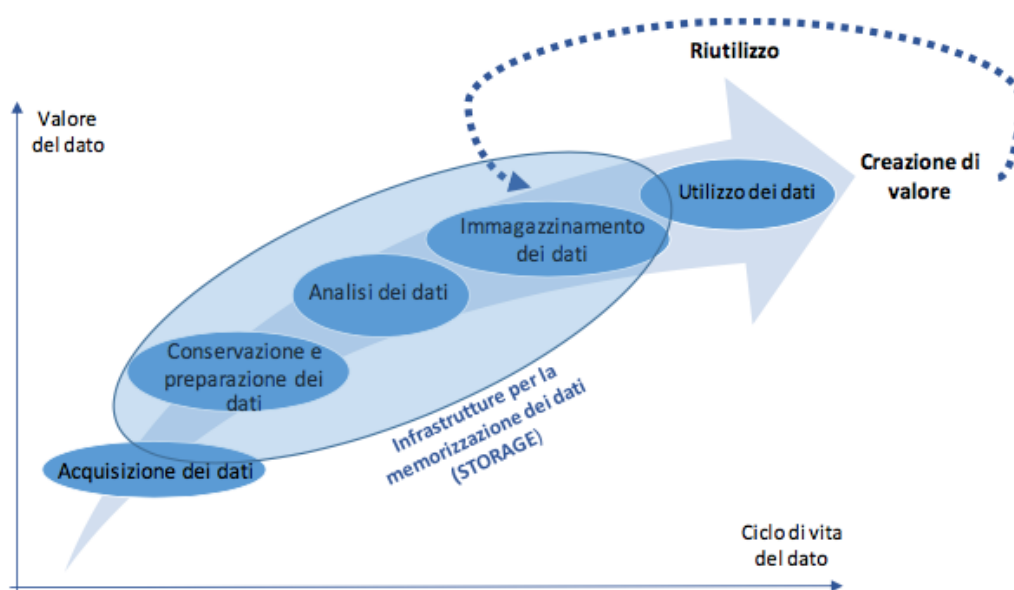


FIGURA 3. La catena del valore nei big data.

(Fonte: Autorità)

Possiamo così distinguere le fasi del processo:

1) Raccolta dei dati. Avviene attraverso una varietà di fonti molto eterogenee tra di loro. In primo luogo, l'utilizzo di telefoni cellulari sempre più innovativi e sofisticati, da parte della maggior parte della popolazione, e, soprattutto, l'utilizzo di Internet, per mezzo del quale si ottengono dati che vengono prodotti sia per accedere a documenti, articoli, brani musicali sia utilizzando motori di ricerca, social network e APP (ovvero quei software applicativi che ciascun utente installa all'interno del proprio device mobile per lo svolgimento di specifiche attività, come ad esempio giochi, editing fotografico, social network, ecc). Inoltre, grande merito è dato anche all'enorme diffusione dell'internet delle cose (IoT) e di sensori di ogni

genere. Con questi mezzi si possono ottenere quantità notevoli di dati che riguardano il profilo personale di ciascuno di noi.

2) Aggregazione, registrazione e conservazione di tutti i dati raccolti. Molte società rendono anonimi i dati prima di registrarli, mentre altre li registrano mantenendo l'identità del soggetto a cui si riferiscono (Richeri, 2016). Inoltre "affinchè i dati grezzi (raw data) si trasformino in informazione, diventa necessario possedere le strumentazioni e le competenze che consentano di affrontare i problemi connessi alla varietà dei dati e, quindi, generare i presupposti affinché i diversi tipi di dati possano interloquire tra loro, spesso in tempo reale" (Agcom, 2018, p. 17). A questo proposito vengono messe a punto infrastrutture immateriali che consentono questo tipo di attività e, soprattutto, l'immagazzinamento dei dati. Queste infrastrutture sono dette *data center*: esse ospitano al loro interno tutte le apparecchiature necessarie per l'elaborazione dei dati.

Importante è sottolineare che gli investimenti in queste infrastrutture stanno aumentando sempre di più poiché le imprese private si stanno rendendo conto della necessità di costruirsi dei propri centri per la raccolta e l'elaborazione dei dati. "Ad esempio, dal 2007 al 2017, Google ha investito circa 3,2 miliardi di euro per costruire i quattro data center che operano attualmente in Europa, per un valore di circa 300 milioni di euro l'anno;" (Agcom, 2018, p. 30)

Inoltre, risulta si stiano diffondendo molto anche i servizi di cloud computing, che costituiscono l'offerta esternalizzata di servizi di immagazzinamento e analisi tramite web.

3) Utilizzo di algoritmi per la correlazione e analisi automatica dei dati raccolti da fonti anche diverse. Da questo momento in poi le attività sui dati sono rivolte al passaggio dallo stadio di semplice di informazione a quello di trasformazione in dato rilevante.

Per l'elaborazione dei big data quindi si utilizzano algoritmi più o meno complessi che servono per comunicare al computer i dati e le elaborazioni da fare con il fine di ottenere le informazioni utili agli obiettivi prefissati.

4) Immagazzinamento dei dati. "La maniera in cui ciascuna organizzazione progetta l'immagazzinamento dei propri file ha un impatto rilevante sulla velocità e sull'efficienza con cui avvengono l'accesso ai dati e, di riflesso, i processi decisionali" (Agcom pag 18)

5) Utilizzo delle informazioni risultanti dalle analisi per la creazione di nuovi contenuti più adatti, per proporre nuove soluzioni agli utenti che corrispondono in maniera più puntuale ai

loro orientamenti di consumo. Questa è la fase finale che fa sì che i dati vengano utilizzati a supporto dei processi decisionali.

1.4 L'industria audiovisiva e il cambiamento rispetto al modello tradizionale

Quelle che maggiormente illustrano e rendono evidente il successo vero e proprio dei big data sono le industrie dei media (o industrie audiovisive).

In particolare nella programmazione televisiva, Netflix, uno dei maggiori colossi (assieme anche ad Amazon), sfida quelli che sono i vecchi canali del tradizionale modello televisivo e vi apporta radicali cambiamenti, per adattarsi alle nuove esigenze e richieste del cliente sempre più esperto e in cerca di novità. Il suo vantaggio competitivo deriva dall'ottenere informazioni dai propri utenti, i quali, secondo il modello di business che Netflix si è stabilito, sono posti al centro delle sue decisioni. Questo rappresenta un importante cambiamento rispetto al vecchio modello televisivo (Fernandez-Manzano,2016). Ora è proprio l'utente, infatti, che potendo accedere tramite un unico canale all'intero catalogo di contenuti, beneficia in prima persona di questo processo.

Internet costituisce così un processo di disintermediazione che ha portato all'emergere di una nuova posizione di dominio attraverso le piattaforme di distribuzione di contenuti online. Queste si basano su una nuova concezione del business dei media, che propone formule incentrate sulla soddisfazione dell'utente

Rispetto al modello tradizionale dell'offerta televisiva, i fattori che differenziano il nuovo modello sono:

- 1) Il fatto che il controllo non è più detenuto da attori direttamente o indirettamente collegati nella produzione e distribuzione dei contenuti
- 2) I circuiti di distribuzione non sono circoscritti in un singolo settore ma si basano sul contenuto che viene distribuito attraverso specifiche piattaforme.
- 3) Ora il pubblico non è più soggetto passivo del processo bensì attivo e ne diventa il principale protagonista, guidando e costituendo il processo di aumento del potere dei contenuti, potendo accedervi liberamente (anche attraverso mezzi non legali) e trasformandolo con le sue interazioni.

Questo nuovo approccio di offerta di contenuti e servizi online, però, non è stato accolto positivamente dalle imprese tradizionali del settore, le quali, ora, si trovano in grosse difficoltà di fronte ai nuovi arrivati.

“Non si tratta, infatti, di un nuovo tipo di televisione derivato dalle competenze di quella tradizionale. Le imprese più innovative e dinamiche, che hanno portato al successo commerciale i nuovi servizi di cui parliamo, vengono da esperienze non televisive e si chiamano Netflix, Amazon, Hulu e altre; alcune imprese leader della tv tradizionale successivamente sono entrate in campo, ma occupano posizioni finora di secondo piano.” (Richeri 2016)

CAPITOLO 2. Analisi del caso Netflix e della sua strategia di business.

2.1 La storia di Netflix: il suo debutto nel mercato televisivo online

Come evidenziato nel precedente capitolo, i *big data* sono fondamentali sia per il processo decisionale, sia per la qualità tecnica del servizio. Forniscono indicazioni su quali siano i migliori contenuti da presentare, come presentarli e a chi presentarli. Il nuovo spettatore richiede una totale libertà quando si tratta di visualizzare contenuti audiovisivi, dove, come e ogni volta che lo desidera.

In questo nuovo settore della distribuzione di contenuti online, Netflix occupa una posizione di rilievo e si distingue per la sua offerta di video, film e telefilm in abbonamento a tutti gli iscritti. Netflix è emersa come azienda di noleggio di video online nel 1997 e il suo modello di business si basa su contenuti che incontrano gli interessi del pubblico e su una nuova e interessante formula monetaria, basata sul canone di abbonamento per il noleggio di DVD. Nel 2007 ha introdotto invece lo *streaming* video, mantenendo il sistema di abbonamento mensile, che ha costituito un punto di svolta nel mercato dei video online. Con la sua innovativa proposta, la società è stata introdotta con successo in diversi paesi europei a partire dal 2012 e continua tuttora a diffondersi senza sosta.

Fino al 2011 è stata la società Apple a guidare il settore; questa infatti aveva come vantaggio l'esperienza nel marketing della musica con il modello di download *pay per unit*, attraverso la piattaforma iTunes. Da quel momento in avanti, l'ingresso di Netflix nel mercato ha sconvolto e cambiato il dominio detenuto dalla Apple, prevalendo su quest'ultima e guidando il settore con un tasso di crescita annuo che si attestava tra il 35% e il 40% (Izquierdo-Castillo, 2015).

Si tratta di “un servizio di streaming dietro abbonamento che, a prezzi più che contenuti (dai 7,99 euro mensili fino a un massimo di 13,99 per poter condividere l'abbonamento con altre tre persone), permette di poter usufruire illimitatamente di un catalogo vastissimo composto da film, serie tv, documentari, programmi di vario tipo del presente e del passato, incluse produzioni originali Netflix.” (Customer Database Analytics, 2018)

I fattori competitivi che finora hanno permesso a Netflix di primeggiare rispetto ai suoi concorrenti più prossimi (come Amazon e Hulu), sono, in primo luogo, il concetto di monetizzazione, ovvero quella formula che permette ai naviganti in rete di accedere a un catalogo molto ampio di contenuti in cambio di una quota di sottoscrizione mensile economica. Un secondo fattore è costituito dalla ricchezza e dalla varietà del catalogo di film, serie tv e documentari che Netflix offre ai suoi abbonati. Vi sono poi altri due elementi fonte di successo per la società: 1) la qualità tecnica delle trasmissioni (che si distingue in SD e HD) e 2) la

mobilità e accessibilità ai contenuti mediante differenti tipologie di dispositivo. Quindi, fondamentali sono l'esperienza dell'utente e la sua soddisfazione rispetto alle proposte che gli vengono fatte. A tale fine Netflix ha sviluppato un complesso software di elaborazione di dati che consiste in un sofisticato e complesso sistema di raccomandazione utilizzato per il disegno, costruzione o ottimizzazione dei contenuti (Izquierdo-Castillo, 2015)

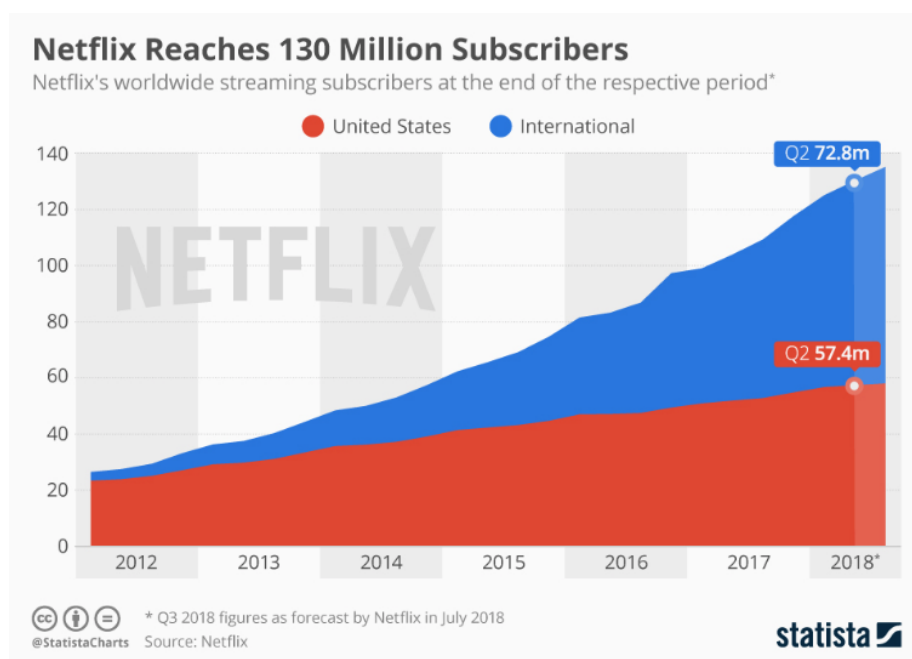


FIGURA 4: Andamento abbonati Netflix fino a luglio 2018
(Fonte: Statista.com)

Alla fine del 2016 Netflix aveva 93,7 milioni di abbonati; nel primo trimestre del 2018 invece, il totale ammontava a 125 milioni (contro poco più di 117 milioni nel precedente trimestre), di cui 57 milioni erano negli Usa e oltre 68 milioni all'estero. “Il colosso del videostreaming ha superato le attese degli analisti con ricavi in aumento del 40% (a 3,7 miliardi di dollari) e utile netto che balza del 63% (a 290 milioni di dollari). Nel primo trimestre dell'anno Netflix ha registrato un aumento record di 7,41 milioni di abbonati al suo servizio in streaming, di cui 5,46 milioni a livello internazionale e 1,96 milioni negli Stati Uniti” (Licata, 2018).

Reed Hastings, co-fondatore, CEO e direttore di Netflix, ha dichiarato che “la recente performance evidenzia una crescita ‘superiore al previsto negli utenti, grazie al contenuto originale e alla continua adozione su scala globale dell'Internet entertainment’, (intrattenimento via Internet). L'azienda intende continuare ad alzare la posta in gioco quando si tratta di investire sul proprio contenuto: ha previsto di spendere tra i 7.5 e gli 8 miliardi di dollari contro i sei miliardi dell'anno scorso.” (Valsania, 2018)

Per l'industria televisiva, l'acquisizione di informazioni dai propri utenti rappresenta un vantaggio competitivo non da poco e il sistema di raccomandazione, come vedremo tra poco, gioca per Netflix, un ruolo fondamentale.

Gli investimenti e le risorse che Netflix totalizza annualmente sono considerevoli. 300 lavoratori si occupano dei cosiddetti 'filtri collaborativi' e 150 milioni di dollari vengono dedicati ogni anno per l'elaborazione di nuovi algoritmi, interfacce per l'utente e piattaforme di raccomandazione (Izquierdo-Castillo, 2015).

Non si tratta, infatti, di un semplice e unico algoritmo, quanto piuttosto di una raccolta di algoritmi diversi che vengono utilizzati per differenti scopi e che si uniscono per creare l'esperienza Netflix completa.

Il catalogo di Netflix si compone di film, serie tv e documentari. Negli ultimi anni, in aggiunta, Netflix ha deciso di investire maggiormente su contenuti di produzione propria, anche se comunque la maggior parte del suo catalogo è composto da contenuti derivanti e ottenuti da concessioni. Queste derivano da accordi raggiunti con Disney (linee Marvel e Pixar), con Fox Television e con CBS Studios. Si tratta però di negoziazioni che aumentano sempre più di complessità quanto più l'azienda cresce e si espande.

La ricerca dei consumatori suggerisce che un utente tipo di Netflix perde interesse dopo circa 60/90 secondi di consultazione, dopo aver esaminato una quindicina di titoli (forse 3 più nel dettaglio) su uno o due schemi. Lo scopo di Netflix, e del suo sistema di raccomandazioni, è quello di far sì che l'utente, nel lasso di tempo in cui effettua la ricerca, trovi con grande probabilità ciò che desidera guardare in quel momento. Se ciò non avviene, aumenta di molto il rischio che l'utente abbandoni il servizio per trovare altro da fare in alternativa. L'obiettivo quindi è assicurarsi che su quegli schemi ogni utente trovi qualcosa di interessante (Gomez, 2015)

2.2.1 Netflix abbandona il feedback a stelline e passa ai giudizi con il pollice di Facebook



FIGURA 5. Un utente che guarda Netflix da un device con connessione a Internet.

(Fonte:Google)

A partire da aprile 2017 è stato introdotto un nuovo sistema mediante il quale gli utenti potevano valutare ciò che guardavano. Prima si valutavano i film attraverso una scala di stelle da 1 a 5. Dopo alcuni test però è stato trovato che un nuovo metodo di valutazione batteva in modo significativo quello precedente basato su stelle. Esso consiste nel valutare i film o i programmi tv con un pollice in su, se piaciuto, o con un pollice in giù se non gradito, ed è affiancato da un nuovo punteggio chiamato ‘percentuale di abbinamento’ per molti titoli (personalizzato per ogni utente), che aiuta gli utenti a trovare più facilmente i video che preferiscono (Simonetta 2017). Si parla di un sistema più intuitivo attraverso il quale si vuole aiutare gli abbonati a capire se un titolo è più compatibile con le loro preferenze. È stato poi riscontrato che effettivamente le valutazioni fornite dagli utenti erano aumentate più del doppio.

Si ritiene che questo risultato derivi da un’idea che degli utenti secondo la quale, fornendo le proprie personali informazioni o giudizi, ottengono in cambio suggerimenti più conformi alle loro aspettative che, quindi, gli permettono di trovare ciò che desiderano più rapidamente (Simonetta, 2017).

2.2 In che modo Netflix raccoglie e analizza i dati dagli utenti. E che dati raccoglie?

“È importante ricordare che il vero valore dei big data non deriva dai dati nella loro forma grezza, ma dalle informazioni che risultano dalla loro elaborazione e analisi. Le innovazioni legate alle tecnologie per i big data e negli approcci gestionali devono essere accompagnate da cambiamenti, altrettanto significativi, nel modo in cui i dati vengono utilizzati per le decisioni aziendali e nella creazione di innovazione in prodotti e servizi.”
(Thomas H. Davenport in Big Data in Big Companies. SAS.com)

“Per capire come lavora Netflix con i *big data*, c'è stato chi ha tentato di ricostruire dall'esterno, attraverso un difficile percorso induttivo, le metodologie di analisi, le descrizioni dei programmi, le combinazioni e i risultati del processo di analisi dei film e degli spettacoli televisivi. Il livello di comprensione raggiunto dopo un lavoro impegnativo e tortuoso è stato però solo un'approssimazione rispetto al sistema assai più utilizzato da Netflix, che si serve di potenti mezzi di trattamento automatico dei dati e di algoritmi appositamente creati, e segreti, per raggiungere risultati efficaci” (Richeri, 2016).

Avere un'ampia base di utenti fa sì che Netflix possa raccogliere un'enorme quantità di dati. Con questi sarà poi in grado di prendere decisioni che consentano di migliorare le tecniche che vengono già utilizzate per rendere gli utenti che utilizzano il servizio più felici e soddisfatti. Le reti televisive tradizionali, invece, non utilizzano questa tecnica ma valutano il da farsi solo basandosi sull'intuito e su ciò che tradizionalmente è stato in grado di soddisfare la clientela. Netflix invece, trattandosi di una società in rete, ha il privilegio di poter sfruttare tutta la mole di dati a sua disposizione, e a disposizione della rete, per approfondire la conoscenza dei propri clienti e personalizzare il servizio. Netflix, ad esempio, è in grado di osservare mediante i dati dei propri utenti, quante persone che hanno ripreso a riguardare la serie interrotta alla stagione 1 sono arrivate fino alla stagione 3. Diciamo che risulti essere circa il 70%. Il resto degli utenti cos'ha fatto? Come mai ha interrotto e non ha continuato? Oppure quanto grande era la differenza temporale tra quando si finiva un episodio e quando si iniziava il successivo? L'obiettivo di Netflix, osservando tutti questi dati, è quello di avere una buona idea su come stanno andando le cose e come migliorarle per far sì che l'utente sia più felice.

Inoltre se la serie è stata poi annullata e conclusa alla stagione 3 ma circa il 70 % delle persone che l'hanno iniziata (ovvero un numero parecchio elevato) sono arrivate fino alla fine, è il caso di chiedersi se valga la pena riprendere il seguito della serie. In tal caso, ci sarebbero grosse

probabilità che un elevato numero di persone seguirebbe l'evolversi della vicenda. Questo, però, è solo un esempio di quello che Netflix studia attraverso i dati che ha a disposizione.

La complessità dei dati raccolti e le informazioni ottenute su ogni abbonato che derivano dall'elaborazione dei dati sono intuibili considerando anche solo alcune tra le attività che gli utenti compiono giornalmente e che vengono registrate automaticamente dal sistema.

Per esempio (Bulygo, 2013):

- il tipo di contenuto scelto (e le caratteristiche principali come protagonisti, regista, genere, epoca ecc)
- quali dei programmi proposti vengono scartati
- il momento in cui si mette in pausa, si riavvolge, o si avanza rapidamente
- in che giorno della settimana si guardano determinati contenuti (è stato provato che le persone che guardano più serie tv durante la settimana e più film durante il fine settimana)
- la data di quando si guarda
- a che ora si guarda
- la località da dove si guarda il programma
- quale dispositivo viene utilizzato
- le valutazioni che si forniscono sui video
- le ricerche che vengono effettuate
- il tempo quotidiano dedicato ai programmi di Netflix.

Inoltre Netflix è in grado di sapere quanti contenuti devono essere guardati da un utente prima che questo decida di annullare l'abbonamento. Ad esempio è stato studiato che se si riesce a far sì che ogni utente guardi almeno 15 ore di contenuti ogni mese, c'è il 75% di probabilità in meno che questo annulli l'iscrizione. Se invece gli utenti guardano meno di 5 ore al mese, c'è il 95% delle probabilità che questo annulli. Una volta scoperto ciò, Netflix ha incominciato a chiedersi quali fossero le giuste mosse per agevolare e incrementare la visione di video da parte degli utenti nel sito. Da qui è nata l'idea dell'opzione che viene chiamata 'post-play' ovvero il momento in cui concluso un episodio di una serie, viene automaticamente in seguito riprodotto quello successivo senza in bisogno che l'utente debba fare nulla. Per quanto riguarda i film invece, quando se ne conclude uno, Netflix propone direttamente sulla schermata del video alcuni suggerimenti di film che potrebbero piacere all'utente, basandosi su quello appena visto, e gli dà la possibilità di fare play direttamente da quella schermata (Fernadéz-Manzano, 2016) In seguito a ciò, e all'utilizzo delle raccomandazioni, è stato verificato un aumento delle ore di streaming, e una riduzione delle percentuali di annullamento dell'iscrizione dal programma.

Secondo quanto riportato da Gomez (2015), gran parte degli annullamenti è dovuto a un errore di pagamento, piuttosto che a una scelta esplicita degli abbonati di abbandonare il servizio.

“Netflix sa bene che se le persone guardano tutti i film a cui sono interessati e poi non riescono a trovare null’altro che stimoli il loro interesse, la conseguenza quasi certa è che l’abbonamento venga cancellato. Per questo motivo utilizza complessi algoritmi in grado di suggerire, sulla base dei programmi visti e di ciò che è piaciuto a chi ha guardato gli stessi programmi, contenuti che hanno un’alta probabilità di essere graditi. Propone insomma prodotti su misura per ogni utente. Il risultato? Circa il 75% delle visioni su Netflix è basata sui suggerimenti: questo significa che il 75% dell’attività dei suoi utenti deriva dalla capacità dell’azienda di sfruttare la customer database analytics.” (Customer Database Analytics, 2018)

Una volta che i dati sono stati preparati e raccolti e sono disponibili in quantità tali da poter essere utilizzati per le esigenze aziendali, il passo successivo da compiere è noto come ‘analisi dei dati raccolti’. In questa fase vengono prese le decisioni su come disporre e modellare la vasta gamma di contenuti in base ai gusti comprovati degli utenti. Inoltre mostra informazioni su come incrementare nuove strategie future e può anche trattare aspetti economici che riguardano per esempio il ritorno sull’investimento (ROI) (Fernandèz-Manzano, 2016, p.573).

2.3 Big Data per l’analisi dei contenuti

2.3.1 Esiste un team di lavoro specializzato

Nel caso di aziende come Netflix, vengono creati diversi *tag* (o ‘microgeneri’) da parte di un gruppo di circa 100 professionisti specializzati e appositamente formati nella visualizzazione e classificazione dei contenuti digitali, che si occupa di scomporre film e programmi televisivi e classificarli considerando ogni sorta di variabile (come per esempio il paese di origine, le caratteristiche fisiche dei protagonisti, le varie componenti della sceneggiatura e della trama, la data di pubblicazione, il regista e moltissime altre variabili). Si tratta di un lavoro di analisi molto sofisticato, che si basa su un documento di 36 pagine (Richeri, 2016) che ogni analista deve seguire per classificare un film o un programma. Tutto viene analizzato al minimo dettaglio. Attraverso questa indicizzazione è possibile identificare fino a 80.000 microgeneri che vanno a riempire il bacino di dati a disposizione dell’azienda per raggiungere l’obiettivo principale prefissato, ovvero di soddisfare al meglio l’utente. Dalle possibili combinazioni di tutte le variabili utilizzate per identificare un film o un qualsiasi spettacolo televisivo, vengono

create centinaia di etichette che possono corrispondere a una molteplicità di film e ogni film può essere catalogato sotto una molteplicità di etichette. Per esempio «Film horror degli anni '80», oppure «Film ambientati nell'epoca vittoriana» e così via. “Basti pensare soltanto ai principali registi, attori e attrici, luoghi, generi e sotto-generi, colonne sonore, epoche storiche, ecc. per individuare un lungo elenco di categorie e loro possibili combinazioni che generano le migliaia di etichette utilizzate da Netflix, destinate a loro volta ad essere combinate automaticamente con gli abbonati che in base al loro profilo possono essere interessati a questo tipo di contenuto.” (Richeri, 2016)

Più informazioni si hanno a disposizione su ogni singolo utente, più saranno accurate le raccomandazioni a lui rivolte e più azzeccati i consigli di acquisto.

2.3.2 *'House of Cards': prima serie originale Netflix di successo*



FIGURA 6. (Fonte:Google)

Fu nel 2011 che Netflix prese una delle decisioni più importanti di sempre, che pose le basi per significanti cambiamenti all'interno dell'organizzazione della società e della sua offerta televisiva. Non si trattava di qualcosa di materiale, bensì si trattava di una decisione riguardante il contenuto. La società, infatti, che aveva già 25 milioni di abbonati negli USA, “decise di investire una bella parte delle sue risorse nella produzione di contenuti originali da offrire agli abbonati in aggiunta a quelli acquistati dalle major hollywoodiane e dalle grandi reti televisive” (Richeri, 2016). Con questa decisione, Netflix otteneva i diritti dalla BBC per creare con i propri mezzi una nuova versione, ambientata a Washington, di House of Cards, serie televisiva tratta da un'opera dello scrittore inglese David Dobbs. Questa, sarebbe stata costituita da due stagioni

divise in 13 episodi ciascuna. Ogni episodio aveva un costo previsto che andava da 4 a 6 milioni di dollari, totalizzando un costo complessivo di circa 100 milioni di dollari (Bulygo, 2013).

Questo fu il primo grande investimento che Netflix fece sul lato dei contenuti. Ma perché ha deciso di investire così tanto in questa missione?

Jonathan Friedland, direttore delle comunicazioni della società, risponde così: “Because we have a direct relationship with consumers, we know what people like to watch and that helps us understand how big the interest is going to be for a given show. It gave us some confidence that we could find an audience for a show like *House of Cards*.”

In un'altra intervista, Steve Swasey, altro dirigente di Netflix, osserva:

“We have a high degree of confidence in [*House of Cards*] based on the director, the producer and the stars.... We don't have to spend millions to get people to tune into this. Through our algorithms, we can determine who might be interested in Kevin Spacey or political drama and say to them ‘You might want to watch this.’”

Swasey continua il discorso affermando che non è solo grazie al cast e al regista che si prevede se lo spettacolo sarà un successo: “We can look at consumer data and see what the appeal is for the director, for the stars, and for similar dramas,” disse. (Bulygo, 2013)

Da ciò si può intuire che quando Netflix decide di produrre una serie, non è necessario che provi ad indovinare cosa il pubblico vorrebbe vedere, poiché ha a disposizione una tale quantità di dati e di indicatori che gli permettono di ridurre notevolmente incertezza e rischio riguardo una possibile risposta da parte del pubblico.

Il caso *House of Cards* rappresenta un chiaro esempio “del potere dei big data associati ad algoritmi di previsione statistica” (Customer Database analytics, 2018). Nel momento in cui bisognava decidere se produrre o no quella serie, Netflix ha applicato quella che viene chiamata ‘business intelligence’ dagli esperti (Fernandèz-Manzano, 2016) e si è arrivati alla sovrapposizione di tre fattori chiave, correlati tra di loro:

- 1) Tantissimi abbonati avevano guardato dall'inizio alla fine il film ‘*The social network*’ di David Fincher
- 2) La versione inglese di *House of Cards* aveva riscosso un grande successo in termine di visualizzazioni
- 3) Coloro che avevano visto la serie *House of Cards*, avevano anche guardato film diretti da Fincher.

“Alla luce di questi tre fattori, e considerando il successo dei thriller a sfondo politico, è venuto spontaneo investire in questa nuova serie con Kevin Spacey e il cui pilot era diretto da Fincher.” (Customer Database Analytics, 2018)

Inoltre, con i dati che avevano a disposizione non si sono ristretti a creare un unico trailer uguale per tutti, ma invece sono stati in grado di creare diversi trailer personalizzati per ogni tipo di membro del sito. Prima della pubblicazione della serie infatti, Netflix ha realizzato 19 diversi trailer (Richeri, 2016) che si orientavano a diverse tipologie di persone, con caratteristiche differenti. Quello che veniva indirizzato a una determinata persona, era basato su ciò che questa aveva guardato in precedenza. Per esempio, a coloro che avevano guardato diversi film con il protagonista della serie Kevin Spacey, veniva proposto un trailer in cui l'immagine dell'attore era preponderante. Diversamente, a coloro che in precedenza avevano guardato molti film nei quali le protagoniste erano principalmente donne, veniva indirizzato un trailer in cui spiccavano le donne della serie, e così via.

Si può dire, quindi, che House of Cards sia stato un successo?

Senza dubbio sì: ha aumentato di 2 milioni di abbonati il bacino di membri di Netflix negli Stati Uniti nel primo trimestre del 2013. Inoltre, ha portato anche 1 milione di nuovi abbonati nel resto del mondo. Secondo The Atlantic Wire, questi 3 milioni di abbonati hanno praticamente coperto quanto investito da Netflix per la serie tv.

E' stato inoltre provato, attraverso diversi sondaggi svolti in periodi successivi all'uscita della serie, che House of Cards rende meno probabile la cancellazione dall'abbonamento per l'86% degli abbonati (Bulygo, 2013).

Esempi come 'House of Cards', 'Arrested Development' e 'Orange is the new Black' sono tutte serie che confermano che le scelte di Netflix su cosa trasmettere, non dipendono mai da fattori casuali, bensì sono frutto di accurate analisi di dati e decisioni prese in base a conoscenze dettagliate sull'ambiente circostante.

C'è anche da dire che la concessione di licenze per la trasmissione di un film o di una serie tv è molto costosa e quindi bisogna capire bene cosa valga la pena fare e cosa no e, in caso, come agire nel migliore dei modi. In questo caso i dati sono il mezzo con il quale Netflix prende le sue decisioni e mostra all'utente ogni giorno nuove proposte di film che potrebbero piacergli.

Quello che richiede la società come contributo tariffario non è abbastanza per potersi permettere di includere qualsiasi tipo di film al catalogo Netflix. Qui entra in gioco l'abilità del team di lavoro per provare a capire quali siano le giuste azioni su cui investire tempo e denaro. Devono prendere decisioni intelligenti e sfruttare appieno le loro analisi per cercare di capire quali sono i film che gli utenti, eventualmente, apprezzerebbero di più. Essere efficienti in termini di costi e rendere felici gli utenti è un'abilità fondamentale per il successo di Netflix.

Chiaramente, però, non è solo attraverso l'analisi dei dati che vengono prese le decisioni; ci sono alcune situazioni in cui entra in gioco anche l'intuizione. Ad un esempio, per la serie tv *Breaking Bad*, i dati non erano in grado di prevedere se lo spettacolo sarebbe stato un successo oppure no. La decisione, in questo caso, come in altri casi, è stata fortemente basata sull'intuito delle persone e della squadra che stava dietro l'idea dello spettacolo. Ciò che l'analisi e i dati possono fare, è fornire informazioni per poter gestire un business migliore e offrire un prodotto superiore. Senza dubbio chi possiede queste informazioni ha un vantaggio superiore rispetto a quelli che agiscono solo per mezzo della propria intuizione o agendo solo secondo ciò che sembra giusto (Richeri, 2016).

2.4 Introduzione ai differenti algoritmi

“La formula segreta di Netflix scorre nel suo algoritmo: si specchia in una home page di titoli che evolve imparando a conoscere cosa ci piace. Arrivando a mettere la foto di questo o quell'attore nella locandina di una serie, se abbiamo apprezzato altri suoi lavori in precedenza; dividendo i programmi in categorie talmente specifiche, che in un dato momento possono comparire anche ad appena lo 0,5 per cento degli iscritti. Una bussola indispensabile per non naufragare in un oceano di migliaia di alternative.” (Morello, 2018)

Il sistema si basa, quindi, su una varietà di algoritmi che costituiscono l'esperienza offerta da Netflix al suo consumatore. La maggior parte di questi si trovano nella home page del sito, ovvero nella prima pagina che un membro di Netflix vede quando vi accede dal suo dispositivo (che sia TV, tablet o telefono cellulare). Qui, infatti, il sistema delle raccomandazioni risulta essere ben evidente ed emergono tutte le caratteristiche su cui esso si fonda.



FIGURA 8: Un esempio di raccomandazioni proposte da Netflix nella Home Page del sito

Nella figura 8 è riportato un esempio della home page di Netflix. Ha un layout a matrice: ogni riga contiene film consigliati che hanno come base lo stesso tema. Questo tema rappresenta l'etichetta di ogni riga e serve a rendere trasparente e più intuitivo il contenuto di ogni riga. Possiamo trovare film nella riga 'Because You Watched', ovvero quei film che ci vengono consigliati per il semplice fatto che ne abbiamo guardati di simili o con alcune caratteristiche che si avvicinavano molto. Un'altra riga di film è invece dedicata a quei film o serie tv o documentari che in un precedente momento avevamo iniziato a guardare, che avevamo momentaneamente interrotto e che con un semplice clic possiamo ricominciare a guardare dal secondo esatto in cui eravamo rimasti.

2.4.1 'Personalized Video Ranker'

I video compresi in una determinata riga provengono in genere da un singolo algoritmo. Ad esempio, tra gli algoritmi noti, l'algoritmo PVR (Personalized Video Ranker) ordina l'intero catalogo di video (o sottoinsiemi selezionati per generi e per altri filtri) per ciascun utente in modo personalizzato ed è specializzato nel compito di assegnare metadati a tutti i contenuti caricati sulla piattaforma (Gomez, 2015). L'ordinamento risultante va a riempire tutte le diverse righe della home page con video selezionati e personalizzati in base alle mosse che il singolo utente ha compiuto in precedenza e alle ricerche che questo ha fatto all'interno del sito. Uno stesso tipo di riga contiene quindi per ciascun utente dei video e dei film differenti e che rispecchiano di più le caratteristiche che gli piacciono dei film che vede (per esempio se il protagonista è uomo o donna, se ci sono animali all'interno della vicenda ecc ecc..).



FIGURA 9: Altri esempi di raccomandazioni nella Home Page di Netflix

2.4.2 'Top-N'

Tra gli altri, c'è anche l'algoritmo Top-N, che include tutti quei film che si pensa possano essere più adatti alle preferenze dell'utente, scegliendo tra le migliori opzioni possibili del catalogo

Netflix. PVR, al contrario, classifica sottoinsiemi arbitrari dal catalogo e non le migliori alternative, come invece fa l'algoritmo Top-N. Quindi il rango Top N comprende algoritmi che guardano solo al capo della classifica del catalogo che l'algoritmo produce, piuttosto che alla classifica per l'intero catalogo. Questi due algoritmi, condividono però anche alcune caratteristiche. Ad esempio, entrambi combinano la personalizzazione con la popolarità e identificano e incorporano le tendenze di visualizzazione su finestre temporali diverse.

2.4.3 *'Trending Now'*

La riga Trending Now (Figura 9 in basso a sinistra) è guidata da un classificatore di tendenza che combina le tendenze temporali a breve termine, le quali costituiscono potenti predittori di video che gli utenti guarderanno, con una buona dose di personalizzazione. Grazie a questo metodo vengono identificati due tipi di trend: quelli che si ripetono annualmente, ma hanno un effetto a breve termine quando si verificano (ad esempio nel periodo di San Valentino vi è un forte aumento della visione di pellicole romantiche), e eventi unici a breve termine (ad esempio l'arrivo imminente di un forte uragano, in un'area densamente popolata, aumenta l'interesse a breve termine di documentari e film su uragani e altri disastri naturali).

2.4.4 *'Because You Watched'*

Un altro tipo di categorizzazione è la riga Because You Watched (BYW). Questa contiene tutti quei film o telefilm che sono associati in qualche maniera o accomunati da qualche caratteristica ai video che l'utente ha guardato in precedenza. L'algoritmo che guida le raccomandazioni in queste righe è detto "sims". Questo algoritmo va a calcolare e associare tutti quei film considerati 'simili' a ogni video del catalogo per poi proporli come opzione di visione a ogni utente, in base ai film che ha già guardato. Anche se quindi la classifica di sims non è personalizzata, va ad essere personalizzata l'offerta di video nella riga BYW della home page a seconda di quali video simili ai precedenti l'utente potrebbe avere piacere di guardare. E nonostante esista una componente commerciale che dà priorità alla visualizzazione di determinati prodotti (come 'Aggiunti di recente'), l'accesso di Netflix alla pagina dei contenuti è un ecosistema dinamico che si adatta alle preferenze di ogni iscritto.

2.4.5 *L'algoritmo di generazione della pagina*

Molto spesso gli utenti hanno stati d'animo differenti da una sessione all'altra, oppure molti account sono condivisi da più di un membro della famiglia. L'algoritmo di generazione della

pagina cerca di risolvere questo problema. Infatti si cerca di offrire una selezione diversificata di file, permettendo così a ogni membro di saltare i video che potrebbero essere una buona scelta per un momento diverso da quello presente, per un'occasione o un membro della famiglia differenti, e identificare rapidamente qualcosa di rilevante nel momento in cui quello stesso utente sta cercando. L'algoritmo di generazione della pagina utilizza l'output di tutti gli altri algoritmi per comporre ogni singola e personalizzata pagina di raccomandazioni, tenendo conto della diversità dei membri e della rilevanza di ogni riga per ognuno.

In pratica, la selezione del contenuto algoritmico rappresenta l'80% delle ore di video riprodotte dagli utenti. Il restante 20% proviene dalla ricerca che ogni singolo utente fa all'interno del sito per esempio inserendo il nome di un attore, di un film o di un genere. Spesso però molte delle parole cercate non sono presenti all'interno del catalogo di Netflix o sono stati scritti concetti troppo in generale, ad esempio parole come 'amore', 'estate', 'famiglia', a ciò Netflix risponde proponendo una serie di opzioni che in qualche modo possono essere collegate a ciò che è stato ricercato. Per esempio inserendo la parola 'amore' verranno proposti tutti quei film o telefilm che nel titolo contengono la parola 'amore', ma non solo. Quindi invece di rispondere alla ricerca non riuscita scrivendo che non è possibile fornire nessun risultato, si prova comunque a offrire qualcosa che rientra in correlazione con quanto inserito, provando ad andare incontro alla richiesta dell'utente.

Quando un utente ricerca una parola o un concetto (anche parziali) vengono attivati una serie di algoritmi. Questi algoritmi di ricerca combinano dati di riproduzione, dati di ricerca e metadati per arrivare ai risultati e alle raccomandazioni che Netflix offre ad ogni utente.

Netflix ha investito molto denaro e molto tempo sugli algoritmi e sui dati. Infatti, si potrebbe pensare che la serie House of Cards sia stata scelta da Netflix solo perché si pensava che agli abbonati potesse piacere la trama. In realtà la questione è molto più complessa di come sembra. La decisione di trasmettere questa nuova serie era basata su una serie di fattori e in gran parte da dati. Netflix infatti si basa principalmente su questi. Analizzare i dati risulta di fondamentale importanza per le aziende, che solo in questa maniera sono in grado di ottenere informazioni dettagliate sui loro clienti, per ottimizzare così i loro processi di marketing e offrire ai propri clienti un prodotto migliore e al passo con i tempi e con le tendenze.

2.4.6 Vecchi e nuovi membri

Il sistema di raccomandazione utilizzato da Netflix compie un lavoro molto più soddisfacente quanto più l'utente è iscritto da maggior tempo al servizio. Dei nuovi membri (ovvero quelli iscritti da meno di 3 mesi), infatti, conosciamo molte meno caratteristiche e preferenze rispetto agli altri. Ad esempio, l'algoritmo PVR tende a classificare i video molto più facilmente e intuitivamente per gli utenti vecchi che per quelli nuovi. Inoltre, ai nuovi membri viene offerto un mese di prova gratis del servizio e, quindi, è molto più facile che in quel periodo siano attivi e presenti all'interno del sito mentre dopo, al momento di dover scegliere se pagare o meno la quota di iscrizione, gran parte di questi abbandona e quindi non permettono al sistema di catturare abbastanza informazioni per adattare e personalizzare il servizio in base ai gusti di ognuno di essi. Mentre invece per gli utenti iscritti da maggior tempo, la percentuale di abbandoni è, rispetto ai nuovi, molto più bassa in quanto loro hanno già provato il servizio da molto tempo e devono semplicemente decidere se pagare o meno un mese in più, su una base di tempo molto maggiore. Pertanto Netflix si concentra molto sulla soddisfazione dei nuovi membri, cercando di trovare nuovi modelli che migliorino le raccomandazioni a loro rivolte e aumentino il loro coinvolgimento (Bulygo, 2013)

2.5 A livello internazionale

Netflix a fine 2015



Netflix nel 2016



FIGURA 10. Presenza mondiale di Netflix negli anni 2015-2016

(Fonte: Netflix.com)

La composizione dei nuovi iscritti, è il segnale che l'espansione internazionale procede a ritmi mai raggiunti fino ad ora e ciò garantisce a Netflix, colosso americano della tv su Internet, lo sviluppo di un business che lo distanzia da tutti i presumibili concorrenti. La società, oggi, ricava ormai la metà delle sue *revenues* all'estero, dove si trova il 55% degli abbonati (Licata, 2018).

Tradizionalmente, i mercati internazionali sono stati un'importante finestra di business per il settore audiovisivo americano. L'Europa, in cifre, costituisce il 40,4% della domanda internazionale per le principali compagnie cinematografiche, e un fatturato di 10.600 miliardi di dollari (Izquierdo-Castillo, 2015). Tutto è reso più semplice e flessibile grazie a uno scenario internazionale sempre più convergente e globale.

Netflix ha iniziato il suo processo di internazionalizzazione nel 2010, con il suo ingresso in Canada. Poi ha diffuso il servizio anche in America Latina e nel frattempo si è preparato a debuttare in Europa nel 2012. Qui, ha iniziato a diffondersi prima di tutto nel Regno Unito e in Irlanda per poi aprire le porte, l'anno successivo, anche a Svezia, Danimarca, Norvegia e Finlandia. Il processo è continuato e, nel 2014, si sono aggiunte alla lista anche Austria, Belgio e Francia (Izquierdo-Castillo, 2015). Nel 2015 è sbarcato anche in Italia.

“I consigli personalizzati che proponiamo agli abbonati sono fondamentali per l'esperienza di Netflix, perché consentono di individuare titoli accattivanti con rapidità e senza sforzo. In linea generale, sono i nostri algoritmi che creano i consigli dopo aver analizzato grandi quantità di dati, ma farli funzionare a livello globale non è stato semplice” spiega Carlos Gomez-Uribe, un produttore degli algoritmi di Netflix. “Durante le nostre ricerche”, spiega Gomez-Uribe, “abbiamo notato che le grandi storie non hanno confini e che gli spettatori di tutte le nazioni hanno in comune molto più di quanto non credano. Ad esempio, uno dei modi usati da Netflix per offrire consigli personalizzati a un singolo utente consiste nell'identificare delle comunità di abbonati con preferenze cinetelevisive simili, quindi proporre all'utente i titoli più amati all'interno di tali comunità. Anziché analizzare il pubblico secondo i criteri di un singolo paese o catalogo, questo sistema individua le comunità globali più affini sulla base dei gusti e delle preferenze personali di un abbonato, poi usa tali informazioni per consigliargli i titoli migliori, indipendentemente dal paese in cui si trova.”

Quindi, le informazioni che giungono da ogni parte del globo, fanno sì che le proposte personalizzate del sito migliorino perché consentono di condividere con gli abbonati i feedback di altri utenti con gusti simili ma che vivono in altre parti del mondo.

Ciò costituisce un vantaggio sia qualora un abbonato dovesse trovarsi in un mercato nuovo e di dimensioni ridotte, sia se si dovesse trovare in un mercato grande e avviato, avendo gusti specifici o di nicchia.

Facendo un esempio, consideriamo le anime giapponesi. Con i dati che Netflix ha in possesso si è identificato un gruppo di abbonati che preferisce il tipo specifico di anime raffigurato nella Figura 9.



FIGURA 11. (Fonte: Netflix Media Center)

“Non sorprende che il paese con il maggior numero di partecipanti a questa comunità sia il Giappone, ma meno del 10% dei membri si trova in Giappone, il resto risiede in altre parti del mondo! In questo caso, aggregare i dati di tutti i paesi ci aiuta a migliorare i consigli per tutti gli abbonati appartenenti al gruppo, a prescindere da dove vivono” commenta sempre Gomez-Urbe (2016).

Oggi i paesi vengono raggruppati in regioni all’interno delle quali i gusti (e quindi anche i cataloghi) sono molto simili e che hanno una base di membri abbastanza grande da generare dati sufficienti per soddisfare tutti i modelli necessari. Vengono quindi eseguite delle copie di tutti gli algoritmi all’interno di ciascuna di queste regioni; ciò fa sì che venga sviluppato un unico sistema di raccomandazione globale che condivide i dati tra i diversi paesi. Netflix cerca quindi approcci che generalizzino molti degli strumenti matematici standard e le tecniche utilizzate per le raccomandazioni per far sì che diversi membri abbiano accesso a cataloghi differenti.

Inoltre si cercano modelli per mezzo dei quali ogni utente possa avere a disposizione, tra le alternative, una lingua di ascolto con cui si sente a suo agio e che conosce per vedere ciò che desidera vedere. Ad esempio se un utente comprende solo la lingua giapponese e si ritiene che a questo utente possa piacere la serie House of Cards, ma non si hanno audio o sottotitoli in giapponese tra le opzioni, il programma non gliela inserisce tra le proposte. Se invece fosse possibile vedere la serie in lingua giapponese, essa andrebbe bene evidenziata a quell’utente.

Come ci spiega Gomez-Urbe (2016): “Part of our mission is to commission original content across the world, license local content from all over the world, and bring this global content to the rest of the world. We would like to showcase the best French drama in Asia, the best Japanese anime in Europe, and so on. It will be too laborious and expensive to cross-translate every title into every other language, thus we need to learn what languages each member understands and reads from the pattern of content that they have watched, and how they have watched it (original audio vs. dub, with or without subtitles), so that we can suggest the proper subset of titles to members based on what they will enjoy.”

Netflix utilizza i dati che raccoglie, quindi, al fine di migliorare la fidelizzazione dei soci a livello globale, ridurre le cancellazioni e ottenere valutazioni positive sul prodotto che offrono. D'altronde, la complessa gestione dei dati cresce ulteriormente mano a mano che altri abbonati si uniscono al servizio e si è arricchita di nuovi livelli conseguentemente all'espansione globale (Fernandèz-Manzano, 2016, p.575). L'aumento dell'efficienza nella progettazione del programma, dovuto a una profonda conoscenza dei propri clienti e alle nuove aperture di mercato, è stato parallelo all'evolversi e all'implementarsi degli Originali Netflix. L'espansione globale ha permesso loro di ridimensionare il proprio pubblico e, grazie ai *big data*, individuare gli stessi segmenti di *viewer* nei diversi territori (Fernandèz-Manzano, 2016, p.575). Le loro nicchie di visualizzazione sono ora globali.

2.6 Una prospettiva *leader-follower*

Le aziende televisive, ma in generale tutte le aziende che stanno investendo enormi capitali per la raccolta e l'utilizzo dei Big Data, stanno generando rendimenti in eccesso e ottenendo vantaggi competitivi che fanno senza dubbio preoccupare tutte quelle aziende che non ne stanno facendo uso e che quindi stanno rischiando di rimanere bloccate nei loro meccanismi tradizionali dai quali potrebbero staccarsi troppo tardi, quando il resto del settore risulterebbe già essere progredito e avanzato a innovazioni sempre più irraggiungibili per competervi dignitosamente.

Secondo un'ottica *leader-follower*, coloro che hanno investito per primi nella raccolta e analisi dei dati, nonostante tendessero a effettuare investimenti molto più ingenti in Big Data, hanno anche generato rendimenti molto più elevati rispetto a quelli dei ritardatari (*followers*). La spesa più elevata, infatti, è correlata a un margine maggiore per la crescita dei ricavi (Ramaswamy, 2013).

Alcuni dei leader del ROI, come Netflix, vedono un grande potenziale nei Big Data per il miglioramento di una serie di attività di marketing, vendita, ricerca e sviluppo e attività di servizi. Questi utilizzano i Big data per identificare nuove opportunità di prodotto e per utilizzare il loro grande potenziale per il miglioramento di quattro attività di marketing: monitoraggio e miglioramento dell'esperienza dei clienti nei canali online (come gli store); monitoraggio delle percezioni esterne del marchio; le mosse dei clienti più esigenti al di là del prezzo; e marketing basato sulla posizione fisica dei clienti (motivo per cui è diventato importante per molte aziende acquistare dati mobili).

I leader sfruttano quindi molto di più, rispetto ai followers, dati non strutturati ed esterni. Nonostante questi siano molto difficili da analizzare per i computer, con l'evolversi delle tecnologie e delle innovazioni per l'elaborazione intensiva di questi dati, tutto ciò sta diventando meno complicato di quello che potrebbe apparire a un primo approccio.

CAPITOLO 3. I rischi legati alla raccolta e all'utilizzo dei Big Data

Nei capitoli precedenti ci siamo concentrati su come vengono raccolte e analizzate le informazioni ricevute dagli utenti e in che misura queste decisioni aziendali vengono prese in considerazione.

“I *big data*, infatti, comportano opportunità e vantaggi significativi, come ad esempio l'acquisizione di nuove conoscenze e la scoperta di nuove relazioni, la personalizzazione delle offerte e delle informazioni, la formulazione di pronostici migliori nei settori più disparati, una maggiore flessibilità delle imprese e delle organizzazioni, l'ottimizzazione di processi e la possibilità di prendere decisioni basandosi maggiormente sui fatti.” (Jarchow, Estermann, 2016)

Nonostante però, i big data presentino delle potenzialità e dei vantaggi significativi, risulta necessario sottolineare che ad essi sono associati anche dei rischi piuttosto gravi che riguardano le informazioni sensibili che vengono messe in circolazione attraverso l'utilizzo di Internet e che rischiano di essere fonte di rilevanti fallimenti di mercato che pregiudicano l'efficienza dell'intero sistema.

“Le questioni più impellenti si pongono nel campo della protezione dei dati personali e nella prevenzione dell'abuso di dati. Alcuni temono una dinamica propria incontrollata dei sistemi dovuta ad algoritmi che nessuno riesce più veramente a capire e a controllare. Un'ulteriore problematica per cui i big data non sono responsabili ma su cui potrebbero avere un effetto potenziatore, è la distribuzione iniqua delle risorse e del potere che rischia di portare a squilibri sociali se non viene controbilanciata per tempo.” (Jarchow, Estermann, pag 1-2)

Il modo in cui sono progettati questi nuovi modelli di business basati sui big data rende spesso difficile implementare tecniche efficaci di conservazione della privacy, in quanto molti dei servizi offerti vengono progettati, proprio come nel caso di Netflix, attraverso il recupero e la raccolta di dati.

In questo capitolo andremo a discutere di questi rischi e proveremo a capire quali sono gli eventuali strumenti che possono venire in aiuto.

3.1 Violazione della Privacy

Tutte le attività che gli individui svolgono navigando su Internet, contribuiscono a alimentare il complesso di dati digitali che sono in circolazione. Sia attraverso la ricerca di informazioni mediante i motori di ricerca, sia attraverso la visualizzazione di film e video, sia con gli acquisti online ecc.. Anche semplicemente connettendo il proprio device l'utente, con il suo indirizzo IP lascia dati relativi alla sua posizione. Una volta che viene creata la scia di dati, non è più possibile per l'utente eliminarla.

Questi dati prima di tutto vengono resi anonimi da chi li raccoglie, poi vengono aggregati e successivamente utilizzati per il processo decisionale aziendale. Per quanto riguarda l'anonimizzazione dei dati, dividere le informazioni generali da quelle che permettono di identificare una persona, è un metodo che 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 e residenza di molti dei soggetti che erano stati resi anonimi (Signore, 2016). A questo proposito, una problematica che si è riscontrata, relativamente ai *big data*, è quella della privacy. Ben presto, infatti, si è riconosciuto il bisogno e la necessità di strumenti tecnici adatti e legali per contrastare i rischi intrinseci di privacy e sicurezza di questi nuovi sistemi di raccolta dati in quanto, le aziende sono pienamente consapevoli del rischio cui si incorre nell'ignorare queste minacce e del fatto che una poco accurata conservazione dei dati o politiche di controllo degli accessi mal progettate potrebbero danneggiare la loro reputazione.

Secondo Brookman e Hans (2013), “molte minacce alla privacy degli individui derivano dal mero atto di raccolta dei dati, indipendentemente dalla successiva lettura di questi dati.”

Tra queste minacce (Fernandez-Manzano, González-Vasco, 2018) vi sono per esempio:

- la violazione dei dati da parte degli hacker. Le conseguenze possono essere misurate sia nel danno finanziario diretto sia in termini di danno morale (poiché vengono rivelate informazioni private e personali di una persona, a volte anche intime o imbarazzanti),
- abuso interno da parte dei dipendenti che, con fini di spionaggio industriale, possono origliare ciò che rivela il cliente,
- perdita del controllo sui propri dati da parte degli utenti una volta che hanno rese pubbliche le proprie informazioni private,

-accesso del Governo senza garanzie legali.

3.1.1 Asimmetrie informative

Gli utenti sono senza dubbio gli attori più vulnerabili in questo processo di raccolta di informazioni che li riguarda in prima persona. Molti di questi, al giorno d'oggi, pensano che il prezzo da pagare per usufruire di determinati servizi sia proprio l'accettazione di politiche sulla privacy e quindi il consenso a trattare i propri dati personali. E per la maggior parte dei casi è proprio così. Prima di poter continuare a navigare o servirsi di un sito web bisogna per forza accettarne le condizioni sulla privacy, altrimenti non è possibile proseguire. Queste politiche però non sono comprese a pieno dalla stra grande maggioranza delle persone che le accetta, spesso senza nemmeno leggerle, e non sono quindi in grado di tenere il conto di quali e quante informazioni vengano acquisite dagli operatori online con cui entrano in contatto e, soprattutto rimangono incerti e ignari dell'utilizzo che di questi dati viene fatto. "Secondo una ricerca condotta dall'Unione Europea nel 2015 (Special Eurobarometer 431 - Data Protection), il 69% degli utenti italiani ritiene la cessione delle informazioni individuali una conseguenza naturale degli attuali stili di vita, e il 54% la mette in connessione con l'accesso ai servizi digitali. Al contempo, il 52% sottolinea come la tendenza a fornire dati individuali non sia priva di rischi" (Agcom, 2018, pag.58)

Lo scambio di informazioni, infatti, "avviene sulla base della presenza di significative e strutturali asimmetrie informative tra gli agenti in causa (in questo caso, tra gli utenti che cedono dati e gli operatori che li acquisiscono e li utilizzano); ciò si tramuta per gli individui in un contesto nel quale, non avendo accesso a tutte le informazioni, è di fatto impossibile procedere a una corretta misurazione dei costi (incerti e potenziali)" (Agcom 2018).

Quindi a prevalere sono gli interessi di coloro che posseggono queste conoscenze, ovvero gli operatori e tutte le aziende che raccolgono big data e li utilizzano. Si può concludere che queste relazioni siano relazioni incomplete e inefficienti poiché manca un vero e proprio meccanismo di mercato.

Da una dimostrazione condotta da un noto avvocato australiano, Maximilian Schrems, si è scoperto "che ci sono grandi aziende, come Facebook, che gestiscono i dati privati degli utenti senza rimuoverli nemmeno quando un utente decide di cancellarsi dalla piattaforma. Diverse azioni da parte dell'avvocato sono riuscite a richiamare l'attenzione delle autorità di controllo

europee che nel 2015 hanno messo in discussione l'accordo siglato in precedenza tra gli Stati Uniti e l'Unione Europea. Successivamente, nel febbraio del 2016 un nuovo accordo, chiamato 'Scudo UE-USA per la privacy', regolamenterà il flusso dei dati per proteggere i cittadini dell'Unione Europea, obbligando le aziende operanti negli Stati Uniti a migliorarne la protezione. Inoltre, con l'ultimo regolamento dell'aprile del 2016, il 'General Data Protection Regulation', si punta a migliorare la sicurezza dei dati personali sia all'interno che all'esterno dei confini dell'Unione Europea, normativa che si pone l'obiettivo di semplificare la gestione delle diverse regolamentazioni vigenti nei diversi stati e di migliorare la protezione dei dati attraverso sanzioni che possono arrivare anche al 4% del fatturato per le aziende che non rispettino le regole stabilite sulla tutela della privacy." (Alfa Group, 2017)

3.1.2. Misure per la protezione dagli abusi

“Allo scopo di evitare abusi e squilibri occorre inasprire le regole di protezione dei dati per l'economia privata, rafforzare il controllo in materia di diritto della concorrenza dei (quasi) monopolisti in Internet e realizzare campagne informative e misure di sensibilizzazione. Inoltre, bisogna migliorare le condizioni quadro giuridiche tese a garantire la sicurezza dei centri di dati e, in concorso con l'economia privata, creare più soluzioni tecniche volte a garantire la protezione dei dati. L'inconveniente di molte di queste misure sta nel fatto che, per sviluppare la loro efficacia, devono essere coordinate su scala internazionale.” (Jarchow, Estermann, pp.1-2)

Bisogna sottolineare anche, però, il fatto che un eccesso di restrizioni farebbe sì che i costi venissero trasferiti dagli individui alle imprese, le quali sarebbero portate ad operare in maniera meno efficiente, “sia da un punto di vista statico (efficienza legata all'allocazione delle risorse), che da quello dinamico (efficienza legata alle innovazioni)” (Agcom, 2018).

3.2 Altri rischi relativi allo scambio di dati

Oltre al problema della privacy, altri fallimenti del mercato spesso si riscontrano perché, in primo luogo, visto che i costi sociali non vengono internalizzati dagli investimenti posti in essere dalle imprese per la raccolta dei dati, si può arrivare a un sovrainvestimento nella raccolta delle informazioni. In secondo luogo perché “in un contesto in cui sono presenti costi di transazione e incertezza riguardo l'assegnazione dei diritti di proprietà sui dati, è probabile che

le forze di mercato non siano in grado di garantire il raggiungimento di una situazione efficiente.” (Agcom, 2018).

Altri rischi sono poi la disinformazione online, l’incompletezza dei contratti che disciplinano i diritti di proprietà sui dati e l’assenza di espliciti mercati che regolino la formazione dei prezzi.

3.1.1 *Discriminazioni*

Come conseguenza dell’utilizzo dei big data, emergono anche possibili pratiche discriminatorie che riguardano principalmente il prezzo. Grazie infatti alle nuove e moderne tecniche che consentono di raccogliere moltissime informazioni sugli utenti, la discriminazione di prezzo diventa ‘perfetta’ e fa sì che avvenga una redistribuzione sociale a danno di specifiche categorie di utenti. Ciò può comportare dei rischi a livello sociale molto significativi se la discriminazione si estende a differenze nella popolazione fondate su etnia, razza o orientamento sessuale.

3.1.2 *Disinformazione e false interpretazioni dei dati*

Nonostante la popolazione attribuisca ai social network un grande rilievo e li consideri come veri e propri strumenti di comunicazione, sono emersi di recente fenomeni di disinformazione online (le cosiddette *fake news*). A ciò contribuisce il fatto che “per mezzo dei social network, i sistemi di personalizzazione automatica (che operano sulla base di algoritmi e dei big data acquisiti), da un lato, e le azioni di condivisione di contenuti informativi compiute dagli utenti, dall’altro, facilitano la proliferazione di notizie false e la propagazione virale di contenuti polarizzanti.” (Agcom, 2018)

Le nuove tecniche per l’elaborazione dei dati si basano infatti su correlazioni che spesso risultano spurie. Il problema sta nel fatto che più dati si hanno a disposizione, e più ampia è la loro varietà, maggiore è la probabilità di trovare variabili che tra loro presentano un legame del tutto casuale. Infatti, se tra due fenomeni ‘è presente una qualche correlazione, non è detto che per forza tra essi vi sia una relazione causale. Da ciò si deduce che basarsi esclusivamente sulle correlazioni è un approccio che può portare a interpretazioni sbagliate.

3.1.3 *Sul caso Netflix*

Sul caso Netflix ci siamo già soffermati e abbiamo ampiamente discusso su come utilizzare i big data a riguardo. In sintesi, è possibile grazie a loro, prevedere con buona approssimazione quali saranno gli orientamenti di consumo dell’abbonato, per far sì che le proposte che gli

vengono fatte lo soddisfino a pieno sui programmi disponibili in catalogo. Il fatto che oltre il 75% degli abbonati segua questi consigli nel momento di scegliere cosa guardare, rappresenta un risultato molto positivo, poiché vuol dire che il trattamento e l'analisi dei big data riescono a prevedere correttamente i loro gusti e quindi riescono a fare proposte azzeccate (Richeri, 2016). Ciò può far sì che la frustrazione che un abbonato è portato a provare al momento della scelta del film da guardare avendone a disposizione una quantità enorme tra cui scegliere, venga ridotta o addirittura eliminata.

“In realtà, è noto che nel mondo della televisione tradizionale, quella «lineare», gran parte delle scelte dei telespettatori è orientata più dall'inerzia che dai loro gusti e interessi. È dunque ipotizzabile che una buona parte degli abbonati a Netflix scelga sulla stessa base dei telespettatori tradizionali, accettando ciò che viene loro consigliato proprio per evitare il «lavoro» o l'impegno che ogni scelta implica.” (Richeri, 2016)

In secondo luogo, l'offerta selezionata di alcuni film per ogni abbonato, può limitare molto il loro consumo televisivo, quando invece avrebbero un intero catalogo di film tra cui scegliere, per scoprire magari generi che, in base ai gusti tradizionali, non si sarebbe mai stati portati a scegliere.

Conclusioni.

➤ Analisi SWOT:

Abbiamo potuto comprendere perchè i *Big Data* costituiscono una risorsa primaria all'interno di un mondo sempre più digitalizzato. Per ricavare da questi il massimo beneficio, è utile riassumere e mettere in luce, attraverso una analisi SWOT, quelli che sono i punti di forza e di debolezza, le minacce e le opportunità che si incontrano durante l'utilizzo di queste nuove metodologie.

Per quanto riguarda i punti di forza, già discussi ampiamente nella prima parte dell'elaborato, uno dei più evidenti risiede nelle 4V: volume, varietà, velocità e valore. “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.”(Signore, 2016). Sarà grazie all'utilizzo dei social network, all'incremento degli acquisti online e al 'IoT' che i *big data* cresceranno sempre di più in tutte le quattro “V”.

Nonostante però oggi vi sia molta più flessibilità tecnica rispetto al passato proprio perchè il campione di dati non è più limitato, bensì molto ampio ed esteso, sono ancora poche le persone che hanno una conoscenza approfondita in materia per capire i dati e per compiere un'analisi approfondita. La carenza di capacità abbastanza elevate potrebbe condurre ad analisi sbagliate e a false associazioni. Inoltre, altro punto di debolezza è la carenza di infrastrutture adatte ad acquisire e memorizzare l'ingente quantità di dati in circolazione, anche se comunque la tecnologia al giorno d'oggi ha fatto grandi passi avanti.

Sulle opportunità c'è da dire che i big data offrono una serie molto ampia di benefici in diversi settori e in diverse aree funzionali che vanno dal marketing, alla contabilità e alla gestione delle risorse umane. Bisogna infatti cercare di sfruttarli al meglio per far sì che i prodotti che le aziende offrono riescano a soddisfare sempre le esigenze dei propri consumatori (Signore, 2016).

In un'ultima analisi, come già discusso nel capitolo 3, abbiamo potuto comprendere che all'aumentare dei dati messi in circolazione, aumenta anche il rischio che la propria privacy venga illegalmente e brutalmente violata da persone esterne. Questa ed altre minacce, come ad esempio l'aumento dei costi per sostenere ritmi molto intensi e gestire quantità elevate di informazioni della popolazione, devono essere monitorate sia da governo e istituzioni, i quali dovranno occuparsi di garantire protezione e sicurezza ai cittadini e evitare che vi siano sfruttamenti illeciti da parte delle organizzazioni più forti sul mercato.

➤ Cosa imparare da Netflix

Una cosa è certa: nessuno di noi, probabilmente, creerà mai un'azienda con lo stesso fatturato di Netflix. Ciò che contraddistingue questo grande colosso americano e che ha fatto sì che l'azienda si differenziasse dalle altre concorrenti nel settore, è il modo in cui si è approcciato al mercato e all'utilizzo che ha fatto dei *big data* riuscendo ad aumentare inverosimilmente il tasso di fidelizzazione e soddisfazione del cliente. L'azienda, infatti, è stata in grado, attraverso l'utilizzo di tecniche di *data mining*, di andare incontro, con le sue proposte, ai gusti degli utenti per offrire un servizio migliore e ciò si è tradotto in una riduzione del numero di cancellazioni dal servizio e in grandi opportunità di business.

Possiamo anche concludere che l'internazionalizzazione del servizio è stato il passo decisivo verso il superamento della frammentazione culturale da parte dei paesi. L'azienda è stata infatti in grado di evolvere nell'ideale di 'nicchia globale' (Fernandéz-Manzano, 2016) e in produzioni di contenuti su misura che rispondessero alle esigenze dei diversi mercati interessati.

La compresenza di questi fattori ha garantito all'azienda ottimi risultati e le previsioni per il futuro di Netflix risultano ancora più soddisfacenti.

Al giorno d'oggi, un numero sempre più considerevole di aziende ha deciso di intraprendere la strada avviata dal pioniere Netflix e da altre in altri settori, poiché i vantaggi che si intravedono sono numerosi e significativi. Inoltre, il percorso per queste aziende *followers* sarà sicuramente più agevole visto che hanno il vantaggio di poter seguire l'organizzazione già strutturata e testata da Netflix. Non bisogna nascondere, però, che, trovandosi di fronte a una vera e propria innovazione, saranno necessari cambiamenti drastici della struttura interna di un'azienda che porteranno a un aumento dei costi di non trascurabile entità che dovranno essere per forza affrontati (sia per quanto riguarda l'implementazione della nuova tecnologia, sia per quanto riguarda la formazione del personale). Il risultato finale però, vedrà l'affermarsi di un vantaggio competitivo di lungo termine, che non sarà ottenuto invece dalle aziende concorrenti che non avranno saputo sfruttare al meglio i *big data*.

Bibliografia

- AGCM/AGCOM, 2018. *Risultati preliminari Indagine Conoscitiva sui Big Data*. Open Gate Italia. Disponibile su: <https://www.opengateitalia.com/themencode-pdf-viewer/?file=https://www.opengateitalia.com/wp-content/uploads/2018/07/Nota-Indagine-Conoscitiva-Big-Data.pdf> [Data ultimo accesso: 03/08/2018]
- AGCOM, 2018. *Big Data*. Interim report nell'ambito dell'indagine conoscitiva di cui alla delibera n. 217/17/CONS. pp. 1-107
- ALFA GROUP, 2017. *Big Data: molti dati ma anche rischi*. Disponibile su: <https://www.alfagroup.it/big-data-molti-dati-ma-anche-rischi/> [Data ultimo accesso: 20/07/2018]
- BIG DATA 4 INNOVATION, 2017. *Dall'IoT al mobile, il data center del futuro nell'era 4.0*. Disponibile su: <https://www.bigdata4innovation.it/big-data/dalliot-al-mobile-data-center-del-futuro-nellera-4-0/> [Data ultimo accesso: 14/07/2018]
- BROOKMAN, J., HANS, G.S., 2013. Why collection matters: surveillance as a de facto privacy harm. *Big data and privacy: making ends meet. The center for Internet and society*. Stanford Law School
- BULYGO, Z., 2013. *How netflix uses Analytics to select movies, create content, and make multimillion decisions*. Neilpatel <https://neilpatel.com/blog/how-netflix-uses-analytics/> [Data ultimo accesso: 18/07/2018]
- BUONFIGLIOLI, F., 2018. *Big data, opportunità e rischi secondo Fosca Giannotti*. Lettera 43. <https://www.lettera43.it/articoli/cronaca/2018/03/31/cambridge-analytica-big-data-fosca-giannotti-sobigdata/219051/> [Data ultimo accesso: 14/07/2018]
- CUSTOMER DATABASE ANALYTICS, 2018. *Imparare da Netflix: fidelizzare i clienti con la customer database analytics*. Disponibile su: <https://data-storytelling.it/customer-database-analytics/imparare-netflix-fidelizzare-clienti-la-customer-database-analytics/> [Data ultimo accesso: 23/08/2018]
- FERNANDEZ-MANZANO, E., GONZALEZ-VASCO, M.I., 2018. *Analytic Surveillance: Big Data business models in the time of privacy awareness. Vigilancia analitica: modelos comerciales de datos masivos y concienciación sobre la privacidad*. El profesional de la información. 27 (2).
- FERNANDEZ-MANZANO, E., NEIRA, E., CLARES-GAVILAN, J., 2016. Data management in audiovisual business: Netflix as a case of study. *El profesional de la información*. 25(4).
- GOMEZ-URIBE, C., 2016. *Un approccio globale per i consigli personalizzati*. Netflix Media Center. Disponibile su: <https://media.netflix.com/it/company-blog/a-global-approach-to-recommendations> [Data ultimo accesso: 23/08/2018]

- IBRAHIM ABAKER TARGIO HASHEM, 2015. The rise of “big data” on cloud computing: Review and open research issues. *Information Systems journal*, 14(C), pp. 98-115.
- IZQUIERDO-CASTILLO, J., 2015. *El nuevo negocio mediatico liderado por Netflix: estudio del modelo y proyeccion en el mercado espanol*. El profesional de la informacion. 24 (6).
- JARCHOW, T., ESTERMANN, B. *Big data: opportunità, rischi e necessità d'intervento della Confederazione. Risultati di uno studio condotto su incarico dell'Ufficio federale delle comunicazioni*. Disponibile su: <https://www.bakom.admin.ch/bakom/it/pagina-iniziale/1-ufcom/informazioni-dell-ufcom/ufcom-infomailing/bakom-infomailing-40/big-data--opportunita--rischi-e-necessita-d-intervento-della-con.html> [Data ultimo accesso: 03/08/2018]
- LICATA, P., 2017. *Big Data e IoT, i passi per trasformare i dati in un vantaggio di business*. Digital4Executive. Disponibile su: <https://www.digital4.biz/executive/big-data-e-iot-i-passi-per-trasformare-i-dati-in-un-vantaggio-di-business/> [Data ultimo accesso: 14/08/2018]
- LICATA, P. , 2018. *Netflix, trimestrale record e boom di abbonati. Ma la crescita è sostenibile?*. Corcom. Disponibile su: <https://www.corrierecomunicazioni.it/media/netflix-trimestrale-record-e-boom-di-abbonati-ma-la-crescita-e-sostenibile/> [Data ultimo accesso: 29/07/2018]
- MORELLO, M., 2018. Netflix, dietro le quinte del gigante dello streaming. *Negli studios della tv globale*. Panorama (15)
- NETFLIX PRICE, 2009. “The Netflix Price”. Retrieved December 6, 2005 from <http://www.netflixprice.com/> [Data ultimo accesso: 20/07/2018]
- NIGHTNGAL, R. (2016), *How Netflix knows exactly what you want to watch*. Makeuseof, <https://www.makeuseof.com/tag/how-netflix-knows-exactly-what-you-want-to-watch/>
- NUNAN, D., DI DOMENICO, M., 2015. Big Data: a normal accident waiting to happen?. *Journal of Business Ethics* 145(3). Pag 481-491
- RAMASWAMY, S., 2013. *What the Companies Winning at Big Data Do Differently*. Harvard Business Review.
- RICCI, G. 2017. “Big Data e valore aggiunto: quali rischi e come affrontarli”. Colin. <http://www.consulentelegaleinformatico.it/2017/02/21/valore-aggiunto-dei-big-data-quali-rischi-affrontarli/> [Data ultimo accesso: 28/07/2018]
- SIGNORE, D., 2016. *Analisi del fenomeno dei Big Data*. Relazione finale, Università LUISS di Roma, Dipartimento di Impresa e Management.

-SIMONETTA, N., 2017. *Netflix abbandona il feedback a stelline. Arriva il pollice, come su Facebook..* Il Sole 24 Ore

-STONE, M.L. (2014), "Big Data for Media", Reuter Institute for the Study of Journalism, University of Oxford, United Kingdom

-VALSANIA, M., 2018. *Netflix:8.3 milioni di nuovi abbonati trimestrali e nuovi record in Borsa.* Il Sole 24 Ore

-WANG, H., et al., 2016. Predicting the Incremental Benefits of Online Information Search for Heterogeneous Consumers. *Decision Sciences*, 47(5), pp 957-981

-XIAOYUAN, S., TAGHI, M.K., 2009. A survey of collaborative filtering techniques. *Hindawi Publishing Corporation, Advances in Artificial Intelligence*. Jun Hong: USA, pp. 1-19