

LUISS



Department
of Business and Management

Chair of Marketing

Dal passaparola ai recommendation systems: una questione di fiducia

Prof. Michele Costabile

SUPERVISOR

Prof. Luca Cascio Rizzo

CO-SUPERVISOR

CANDIDATE

Francesco Lo Storto - 227511

Academic Year 2020/2021

*Ai miei genitori,
Mia guida ed esempio quotidiano
A Lele e a Giulia,
Spinta imprescindibile di ogni mio sogno*

SOMMARIO

La rivoluzione digitale ha comportato numerosi cambiamenti nelle nostre vite, ed uno dei più impattanti riguarda i canali e gli strumenti attraverso cui passa la comunicazione. Nel mondo delle imprese e nel rapporto che queste hanno con i consumatori è stato possibile osservare come il cambiamento abbia colpito il modo in cui i consumatori stessi si influenzano tra loro durante il processo di acquisto. Difatti si è potuto osservare il passaggio da una forma di passaparola che potremmo definire “di prossimità”, intendendo con questa espressione il Word of Mouth tradizionale, ovvero il passaggio di informazioni e giudizi riguardo un prodotto o un servizio tipicamente tra conoscenti e attraverso canali diretti, ad una forma di passaparola elettronico, che potremmo definire “on-line”, nel quale sono cambiati sia i rapporti tra mittente e destinatario, che il canale, tipicamente indiretto. Per le imprese però risulta complicato inserirsi in questo continuo passaggio di informazioni che ha luogo sul Web, anche perché oltre alle difficoltà di controllo, come verrà analizzato nel primo capitolo, un passaparola guidato dall’impresa gode di una scarsa fiducia da parte del consumatore, rendendo le informazioni poco credibili. Per tale motivo questo lavoro ha lo scopo di offrire una via alternativa per i brand per avere maggior controllo nel processo decisionale del consumatore, e questa nuova strada è fondata sui *recommendation systems*, i quali attraverso i dati raccolti, analizzati ed elaborati, offrono la possibilità di guidare le scelte del cliente attraverso la personalizzazione di massa dell’offerta digitale.

INDICE

Introduzione

1. WOM ed e-WOM

1.1 WOM

1.2 e-WOM

1.2.1 La digitalizzazione, l'acceleratore del passaparola

1.2.2 Gli elementi dell'e-WOM, e la loro influenza sull'intenzione di acquisto

1.3 WOM, e-WOM e viral marketing

1.4 L'importanza del tipo di linguaggio adottato dal comunicatore

2. Recommendation systems: il passaparola in un algoritmo

2.1 Da brick a click

2.2 La tecnologia a supporto dell'impresa

2.2.1 Big data

2.2.2 AI e Machine Learning

2.3 Recommendation systems

2.3.1 Vantaggi e svantaggi dei recommendation systems

2.3.1.1 Pro e contro nei sistemi *content based* e *collaborative filtering*

3. Case studies

3.1 Netflix

3.1.1 L'alba dei sistemi di raccomandazione

3.1.2 Gli algoritmi alla base dei suggerimenti di Netflix

3.2 Zalando

3.2.1 I modelli di raccomandazione utilizzati da Zalando

3.2.2 L'architettura del sistema

3.3 Sviluppo di un pensiero critico personale

ELENCO DELLE FIGURE

1. Affidabilità percepita dai consumatori dei canali di comunicazione
2. Fasi processo decisionale
3. Percezione dei contatti con social network
4. Framework teorico per integrare la tecnologia nell'impresa
5. AI nelle imprese
6. Schema sistema di raccomandazione proposto da Lun-ping Hung
7. Clusterizzazione dei consumatori, dei prodotti e proposte uniche di vendita 1
8. Clusterizzazione dei consumatori, dei prodotti e proposte uniche di vendita 2
9. Clusterizzazione dei consumatori, dei prodotti e proposte uniche di vendita 3
10. I 6 step del modello di raccomandazione
11. Netflix homepage 1
12. Netflix homepage 2
13. Long-Short Term Memory

ELENCO DELLE TABELLE

1. Punti di contatto e-WOM con fasi del processo decisionale
2. Risultati sondaggio Noraini, Kanyan e Nazrin
3. Variabili in cui WOM ed e-WOM differiscono
4. Imprese data driven in Europa

INTRODUZIONE

Come afferma Rasetti¹ lo sviluppo dell'IoT, che porterà in meno di dieci anni ad avere ben più di 150 miliardi di sensori connessi in rete, sta aumentando in modo esponenziale la produzione di dati, che ogni anno raddoppia in quantità rispetto all'anno precedente, e per comprendere la portata di questo evento, basti notare come nel 2018 sono stati prodotti tanti dati quanti ne erano stati prodotti dall'umanità fino al 2017, cioè dai primi segni scalfiti sulle mura delle grotte all'alba dell'umanità, passando per le pergamene egizie, le opere di Omero e Tacito, fino alle e-mail e gli articoli di giornale redatti negli anni più recenti. La centralità dei dati nell'economia moderna risulta dal fatto che questi rappresentano una risorsa preziosa per le imprese per analizzare e soddisfare al meglio il bisogno del consumatore attraverso forme di comunicazione alternative a quelle tradizionali, come i *recommendation systems*. Difatti poichè la comunicazione tra consumatori, definita Word of Mouth, che con l'avvento di Internet e del *marketspace* digitale si è evoluta nell'Electronic Word of Mouth, può rappresentare un potenziale pericolo per le imprese, non avendo queste il controllo su tale forma di comunicazione, una risorsa alternativa è rappresentata proprio dai sistemi di raccomandazione. Questi difatti, raccogliendo ed analizzando la digital *footpring*, ovvero l'ingente mole di dati che ogni individuo genera sul Web con le sue azioni, ed analizzandoli con complessi algoritmi, riescono ad implementare una raccomandazione personalizzata per ogni consumatore che si traduce in una personalizzazione di massa dell'offerta attraverso le homepage ed i layout dei siti di e-commerce o delle piattaforme online. Il lavoro si divide in tre capitoli attraverso i quali verranno analizzati i temi appena introdotti, avendo come filo conduttore la fiducia dei consumatori, fattore dal quale nessuna impresa può prescindere senza vedere una riduzione drastica dei propri ricavi e della sua brand image. Nel primo capitolo vengono analizzati i tratti principali di WOM ed e-WOM, sottolineando come la digital transformation abbia guidato un cambiamento nelle modalità con cui i consumatori comunicano tra loro, studiando quali sono gli elementi del passaparola digitale che influenzano le scelte del consumatore con un focus anche sul ruolo del linguaggio adottato dal comunicatore. Il secondo capitolo si sviluppa tenendo in considerazione le opportunità che le nuove tecnologie, come l'intelligenza artificiale e il Machine Learning, riescono ad offrire alle imprese, quali la possibilità di generare, grazie all'analisi dei Big Data, una forma di passaparola internalizzato e se possibile quanto meno percepibile in modo diretto dal consumatore quale il *recommendation system*, di cui a fine capitolo vengono analizzate le principali forme con due teorie applicative. A conclusione il terzo capitolo ha come oggetto lo studio di due imprese quali Netflix e Zalando che fanno dei recommendation systems il motore del loro successo.

¹ M.Rasetti, "Il dado è tratto. Big data e IA tra scienza e società"

CAPITOLO 1

1.1 WOM

Il passaparola rappresenta la capacità di un marchio di far parlare di sé, attraverso una comunicazione di marketing indiretta, in modo virale, una contaminazione informativa che non conosce confini spazio-temporali, il “*modo usato per trasmettere rapidamente e sottovoce un ordine da una estremità all'altra [...], consistente nel farlo ripetere da ciascuno al vicino successivo*”². Ma addentrandoci nello specifico, il WOM ha avuto diverse definizioni da parte degli studiosi, come quelle di Arndt, Westbrook e Barreto³, in cui il minimo comune denominatore è dato dalla definizione del passaparola come una comunicazione informale tra soggetti, riguardante un oggetto o una esperienza con lo scopo di condividere informazioni ma con un intento non commerciale. Da una analisi effettuata dalla società che si occupa di ricerche di mercato Nielsen⁴ riguardo ai mezzi di comunicazione che più influenzano i consumatori, come mostrato dalla Figura 1, nel 2015 in Italia il passaparola tradizionale, ovvero il passaggio informativo tra conoscenti, rappresenta la forma promozionale più credibile con il 74% di persone che affermano di fidarsi di questa forma di comunicazione attendibile, e subito a seguire con il 64% ci sono le *review* online: dunque WOM ed e-WOM, in termini di fiducia del consumatore, dominano nel panorama delle diverse tipologie di pubblicità. Ovviamente ci sono diversi livelli di necessità informativa da parte dell'individuo, e facendo ricorso alla classificazione dei prodotti proposta da Nelson, in base alla possibilità di conoscerne l'effettiva qualità degli stessi, possiamo descrivere tre macro categorie: i beni di ricerca, per i quali il consumatore non necessita di un elevato supporto informativo, conoscendo le qualità anche prima dell'acquisto, i beni di esperienza, per i quali è necessario per il consumatore acquistare e utilizzare il prodotto o il servizio per giungere alla conoscenza delle caratteristiche, ed infine i beni di fiducia, per i quali difficilmente anche dopo l'utilizzo il cliente riuscirà a capirne a fondo la qualità. Tenendo conto di questa suddivisione, Harrison-Walker⁵ afferma che quanto più un prodotto è di difficile valutazione prima dell'acquisto, classificandosi dunque come bene di esperienza o in modo ancor più marcato in bene di fiducia, tanto più saranno determinanti nel processo di acquisto le informazioni ottenute dagli altri

² <https://www.treccani.it/vocabolario/passaparola>

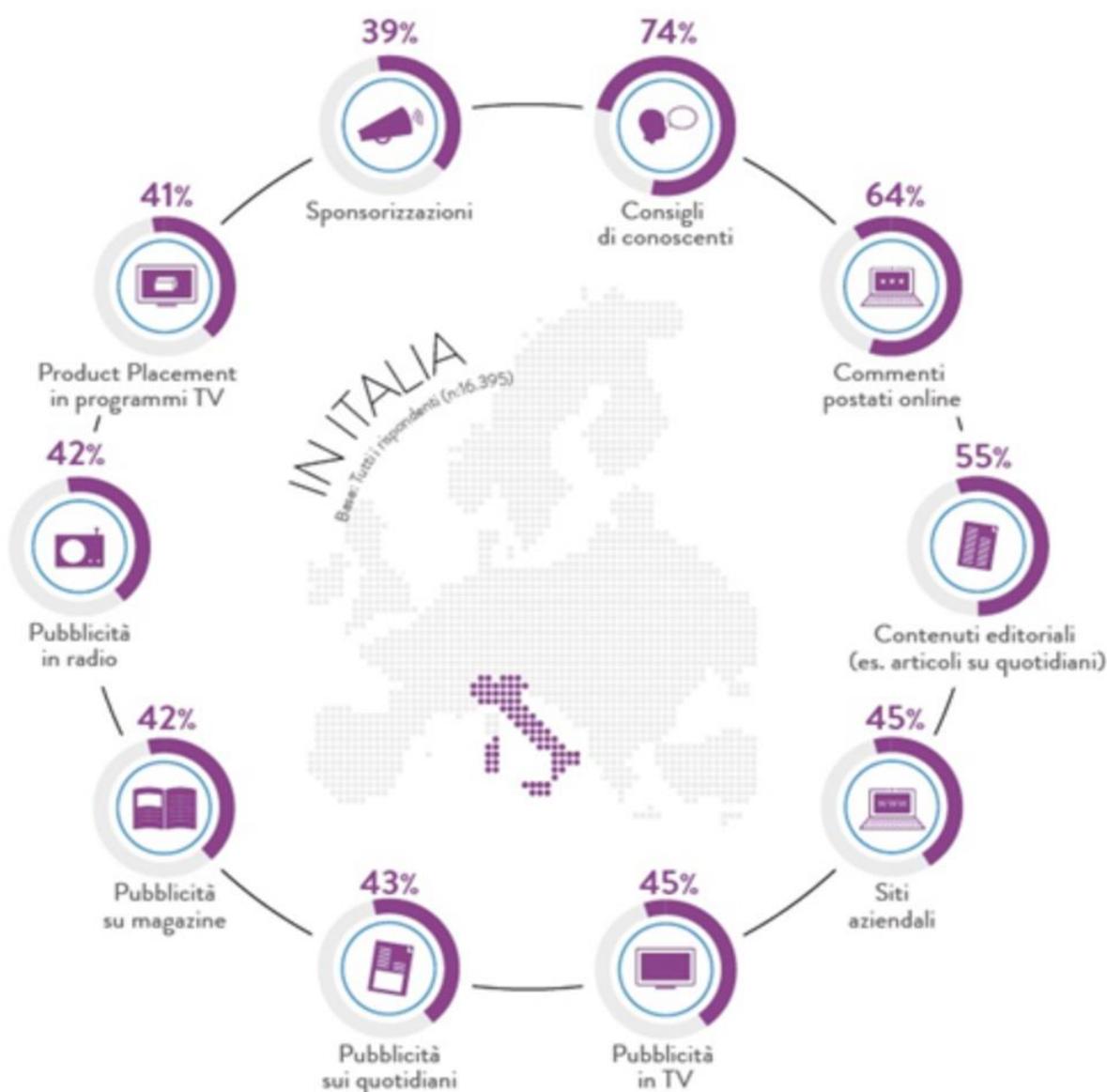
³ Past, Present, and Future of Electronic Word of Mouth (EWOM), Sanjeev Verma & Neha Yadav

⁴ <https://www.nielsen.com/it/it/insights/report/2015/trust-in-advertising/>

⁵ The Effect of E-WOM on Customer Purchase Intention, Noraini Sa'ait, Agnes Kanyan and Mohamad Fitri Nazrin

consumatori, fattore che risulta determinante in misura maggiore per i servizi, data la loro intangibilità. Questo ultimo punto è confermato dai dati che mostrano come il 67% dei cittadini statunitensi trovano un medico grazie al passaparola di amici e parenti, ed in un sondaggio del 2014 di ProCusWright, il 75% dei turisti ha affermato di prenotare un hotel dopo aver consultato TripAdvisor mentre, dato ancora più significativo del ruolo cruciale che il WOM ha nel processo di acquisto, e della fiducia che i consumatori hanno nel giudizio delle persone che hanno fatto esperienza di un servizio, il 53% ha affermato di non prenotare un hotel se privo di recensioni.

Figura 1: Affidabilità percepita dai consumatori dei canali di comunicazione



1.2 e-WoM

1.2.1 La digitalizzazione: acceleratore del passaparola

Lo sviluppo di Internet e la digitalizzazione hanno accelerato in modo esponenziale le informazioni che si rincorrono tra i consumatori, di cui questi hanno urgente bisogno per superare le due principali barriere che si frappongono tra loro e la scelta di acquisto, ovvero il rischio e l'incertezza. Da parte delle imprese lo sviluppo di canali completamente digitali, nei quali hanno luogo le interazioni tra queste e i loro clienti, ha richiesto l'aggiunta all'analisi del marketplace, quella del marketspace, ovvero non più il *consumer behavior* tra gli scaffali degli *store* fisici, ma anche tra le righe di codice dei siti degli *store* online. Per sottolineare il sempre più elevato numero di Internet user, alcuni studiosi parlano di Word of Mouse e non più di Word of Mouth, ricalcando il pensiero di Gupta e Kim⁶ secondo i quali il Web può essere rappresentato come un coffe shop dove le persone con interessi affini possono trovarsi ed iniziare il loro scambio di informazioni elettroniche, avendo la possibilità attraverso il Web 2.0 di stabilire una comunicazione *one-to-one*, *one-to-many* e *many-to-many* tra i consumatori o tra questi e le imprese.

Dai dati ottenuti dallo studio annuale di We are social⁷, è possibile osservare una esponenziale accelerazione della digital transformation, e in particolare in questo caso della presenza sul Web della popolazione mondiale, in continua crescita, con un aumento dal Gennaio 2020 di 93 milioni di persone per l'uso degli smartphone, incremento dell'1,8%, di 316 milioni di internet user, più 7,3% sull'anno appena concluso e di 490 milioni di active social media user, aumento del 13,2%. Se andiamo ancora più a fondo nell'analisi di questi dati, bisogna sottolineare che il tempo medio giornaliero speso da ogni user utilizzando internet è di ben 6 ore e 54 minuti, cioè un terzo della nostra giornata, mentre il 92,6% degli utenti accede al Web tramite il proprio smartphone, mostrando la schiacciante prevalenza di questo device, diventato oramai una vera e propria estensione del nostro corpo, degenerando purtroppo sempre più spesso nel fenomeno della nomofobia⁸, ovvero della dipendenza ossessiva dal proprio telefono, come sottolineato da uno studio della City University di Honk Kong.

Questa trasformazione dei canali di contatto ha portato le imprese ad intravedere la possibilità di aumentare la percezione del valore e la brand image del consumatore attraverso lo sviluppo dell'analisi dei *consumer insight* con tecniche come *il text mining*, *la sentiment analysis*, *l'hashtag analytics* e molti altri strumenti che traggono informazioni dalla ingente mole di dati prodotta dai

⁶ Past, Present, and Future of Electronic Word of Mouth (EWOM), Sanjeev Verma & Neha Yadav

⁷ <https://wearesocial.com/digital-2021>

⁸ <https://tech.everyeye.it/notizie/nomofobia-fenomeno-in-crescita-smartphone-diventa-estensione-del-corpo-301934.html>

consumatori attraverso l'e-WOM, lasciando ampio spazio per la creazione di algoritmi di raccomandazione che hanno lo scopo di personalizzare l'offerta rendendola il più aderente possibile alle esigenze di ogni individuo. Data la centralità che ha assunto il passaparola tradizionale, ed in misura ogni anno maggiore la sua forma elettronica, analizziamo questo fenomeno partendo dalla teoria di Hovland⁹, uno dei padri fondatori della ricerca nel campo della comunicazione sociale, secondo il quale l'analisi di WOM ed e-WOM, si può riassumere in un contesto in cui il comunicatore, anche definito *brand evangelist* da Kawasaki¹⁰, nel nostro caso un cliente che ha già fatto esperienza del prodotto o del servizio o l'impresa stessa, trasferisce uno stimolo, attraverso il passaggio di informazioni, che può modificare il comportamento del ricevente, nel nostro caso la scelta di acquisto da parte di un potenziale cliente, tenendo conto però del rumore che nella teoria tradizionale della comunicazione risulta essere l'insieme di tutti quei fattori che possono distorcere il messaggio durante il passaggio da comunicatore a ricevente, che nel caso dell'e-WOM può essere rappresentato dalla piattaforma sulla quale avviene il passaggio informativo. A riguardo Chatterjee¹¹ afferma che nel processo di scambio di informazioni c'è un fattore chiave che rende i dati acquisiti da questo fenomeno una fonte sicura per le imprese riguardo l'andamento delle preferenze dei consumatori, rappresentato proprio dal continuo scambio di informazioni che prescinde da qualsiasi influenza dei brand. La chiave di lettura si sostanzia dunque nel fatto che gli acquirenti credono più ai pareri di chi ha già avuto esperienza del prodotto o del servizio oggetto della decisione di acquisto, piuttosto che ai consigli che le imprese stesse rivolgono attraverso elaborati algoritmi che hanno assunto elevata importanza attraverso i *recommendation systems*, tema che analizzeremo in modo approfondito nei prossimi capitoli.

Oltre a quelle appena analizzate, la letteratura offre diverse teorie riguardo il passaparola elettronico, descritte da Mishra e Satish¹² per avere un quadro completo del background concettuale alla base dell'e-WOM. Tra queste teorie possiamo annoverare il modello di adozione dell'informazione con il contributo di Sussman e Siegal¹³, dal quale viene studiato il comportamento che i consumatori

⁹ The impact of electronic word-of-mouth communication: A literature analysis and integrative model, Christy M.K. Cheung, Dimple R. Thadani

¹⁰ The differential impact of brand loyalty on traditional and online word of mouth: The moderating roles of self-brand connection and the desire to help the brand, Jiska Eelen, Peren Ozturan, Peeter W.J. Verlegh

¹¹ The Effect of E-WOM on Customer Purchase Intention, Noraini Sa'ait, Agnes Kanyan and Mohamad Fitri Nazrin

¹² eWOM: Extant Research Review and Future Research Avenues, Anubhav Mishra and Satish

¹³ eWOM: Extant Research Review and Future Research Avenues, Anubhav Mishra and Satish

possono avere dopo la ricezione di una informazione, il modello di probabilità di elaborazione che invece analizza come le caratteristiche di credibilità e valore del comunicatore influiscono sulla elaborazione dell'informazione, notando come ci siano due modi per influenzare un individuo, ovvero con un segnale diretto o indiretto, a seconda che ci sia un alto coinvolgimento del consumatore o viceversa, laddove, collegandoci alla teoria di adattamento cognitivo, per coinvolgimento si intende la motivazione e la capacità di elaborare l'informazione da parte del consumatore. Un'altra teoria alla base dell'analisi dell'e-WOM è il modello dello scambio sociale, che affronta le motivazioni per cui un consumatore è portato a condividere al pubblico del Web la sua esperienza d'acquisto e di utilizzo di un prodotto, mentre invece la teoria del flusso multi step si concentra sulle diverse fasi che portano una informazione ad una diffusione virale, soffermandosi sul ruolo del leader di opinione, ovvero colui che è in grado di avere una influenza maggiore degli altri consumatori attraverso i suoi pareri, figura oggi riscontrabile negli influencer. Infine, la teoria della aspettativa e della conferma di Oliver¹⁴ sottolinea che il consumatore sarà soddisfatto o meno di un prodotto a seconda del valore di aspettativa e della conseguente performance del prodotto stesso.

1.2.2 Gli elementi dell'e-WOM, e la loro influenza sull'intenzione di acquisto

Dopo aver osservato diverse teorie alla base del passaparola elettronico, è utile focalizzare l'attenzione sugli elementi che lo costituiscono, approfondendo l'influenza che questi hanno sull'intenzione di acquisto. Prima di analizzare gli elementi dell'e-WOM è necessario però contestualizzare il discorso riguardo l'intenzione di acquisto: questa difatti fa parte del processo decisionale di acquisto, che a sua volta Dewey¹⁵ ha diviso in cinque fasi schematizzate e descritte nella Tabella 1 sottostante, per ognuna delle quali viene fornito un esempio di contatto con il passaparola elettronico, e più in generale con forme di informazione digitale.

Tabella 1: Punti di contatto e-WOM con fasi del processo decisionale

FASE DEL PROCESSO DECISIONALE	ESEMPI DI PUNTI DI CONTATTO CON e-WOM
Riconoscimento del bisogno	Stimoli esterni come pubblicità targetizzate con sistemi di raccomandazione
Ricerca dell'informazione	Motori di ricerca, siti di e-commerce, blog, social network

¹⁴ eWOM: Extant Research Review and Future Research Avenues, Anubhav Mishra and Satish

¹⁵ eWOM: Extant Research Review and Future Research Avenues, Anubhav Mishra and Satish

Valutazione delle alternative	Analisi comparata di recensioni, video, Social network, consigli di influencer (leader di opinione dell'acquirente)
Intenzione e decisione di acquisto	Acquisto su e-commerce o tramite social network
Comportamento post acquisto	Fidelizzazione mostrata tramite recensioni positive, condivisioni sui social network, giudizi e consigli sui blog

Un riscontro della centralità del passaparola nelle varie fasi del processo decisionale ci è offerto dal sondaggio condotto da Cheung e Thadani¹⁶, il quale ha evidenziato come il 91% dei consumatori consultano le *review* durante il processo di acquisto online, e tra questi il 46% ha affermato di esserne influenzato. Tornando all'approfondimento oggetto di questo paragrafo, focalizziamo la nostra attenzione sull'analisi degli elementi che costituiscono l'e-WOM, ed in particolar modo partendo dalla teoria di Noraini, Kanyan e Nazrin¹⁷, che delinea una suddivisione del passaparola elettronico in quattro elementi fondamentali: la rilevanza, la accuratezza, la tempestività e la completezza. Il primo elemento risulta fondamentale perché, come sottolineato da Madu¹⁸, nel processo di ricerca che i consumatori compiono con lo scopo di carpire informazioni utili su un prodotto o un servizio, il processo di lettura di una pagina di un blog, non avviene in modo dettagliato e sequenziale, bensì attraverso una scannerizzazione della pagina alla ricerca delle informazioni di cui si ha bisogno, per l'appunto le più rilevanti, cercando di massimizzare la velocità minimizzando lo sforzo. Per questo motivo quanto più una community, un blog o una sezione di recensioni presenta una selezione accurata, tanto più questa avrà successo nell'e-WOM. Il secondo elemento del passaparola elettronico, la accuratezza, che può essere strettamente interconnesso alla credibilità, fattore trattato nel prossimo paragrafo, e alla qualità della fonte, risulta essere un fattore che aumenta in modo esponenziale la percezione della fonte da parte del ricevente. L'asincronia inoltre rappresenta un punto di forza dell'e-WOM, ma può senz'altro celare un rischio di ininfluenza qualora la ricerca di informazioni da parte del consumatore riscontri recensioni non aggiornate, le quali non aggiungono

¹⁶ The impact of electronic word-of-mouth communication: A literature analysis and integrative model, Christy M.K. Cheung, Dimple R. Thadani

¹⁷ The Effect of E-WOM on Customer Purchase Intention, Noraini Sa'ait, Agnes Kanyan and Mohamad Fitri Nazrin

¹⁸ The Effect of E-WOM on Customer Purchase Intention, Noraini Sa'ait, Agnes Kanyan and Mohamad Fitri Nazrin

nessun valore al processo decisionale di acquisto, portandoci dunque a sottolineare l'importanza della tempestività dell'informazione o del suo aggiornamento nel contesto asincrono del Web. Infine, la completezza del messaggio è direttamente proporzionata alla consapevolezza del prodotto acquisita dal consumatore, migliorando così la probabilità di acquisizione e ritenzione del consumatore.

Con una analisi di regressione multipla notiamo in che misura i quattro elementi appena analizzati sono relazionati con l'intenzione di acquisto dei consumatori: i dati acquisiti da Noraini, Kanyan e Nazrin, mediante un questionario somministrato a 361 individui selezionati dalla popolazione della Malesia, sono stati divisi in tre sezioni, con la sezione A con dati anagrafici, la sezione B nella quale è stata richiesta l'influenza che ogni elemento dell'e-WOM ha nella intenzione di acquisto, avendo come possibili risposte cinque livelli della scala Likert, dati da fortemente d'accordo, d'accordo, neutrale, in disaccordo e fortemente in disaccordo, e la sezione C nella quale è stata misurata l'intenzione di acquisto dell'individuo usando anche in questo caso i cinque livelli della scala Likert. Da questi dati è stata poi effettuata una analisi di correlazione tra i quattro elementi e l'intenzione di acquisto per capire l'influenza che la rilevanza, la accuratezza, la tempestività e la completezza hanno sul consumatore, avendo i risultati riportati in Tabella 2, dalla quale notiamo come secondo questo sondaggio l'accuratezza e la completezza del passaparola sono gli elementi con il più alto grado di influenza sull'intenzione di acquisto del consumatore.

Tabella 2: Risultati sondaggio Noraini, Kanyan e Nazrin

	Rilevanza	Accuratezza	Tempestività	Completezza
Intenzione di acquisto	0,605	0,759	0,565	0,688
N campione	361	361	361	361

Una ulteriore analisi svolta da Lòpez e Sicilia¹⁹ ha portato alla luce altre quattro determinanti dell'e-WOM che influenzano il consumatore durante il processo di acquisto: tra queste troviamo tre elementi del processo di comunicazione, rispettivamente la percezione della valenza, del volume e il tipo di sito tramite il quale avviene il passaparola, e un quarto elemento legato ad una specifica qualità del ricevente, ovvero la sua esperienza con il Web in termini di abilità nell'uso e di tempo già passato su Internet. Analizzando i primi tre elementi, la percezione della valenza risulta essere strettamente

¹⁹ Determinants of E-WOM Influence: The Role of Consumers' Internet Experience, Manuela Lòpez and María Sicilia

legata alla impossibilità da parte di un individuo di leggere attentamente ogni recensione presente su un blog o sul sito di e-commerce di un brand, con la conseguente ricerca delle informazioni che sono appunto percepite come più rilevanti. Su questo punto è necessario evidenziare differenti teorie riguardo all'importanza delle recensioni negative rispetto a quelle positive: difatti da un lato Park e Lee sottolineano la netta superiorità, in termini di peso specifico della valenza percepita, delle informazioni negative rispetto a quelle neutre o positive, portandoci a distinguere la valenza oggettiva da quella soggettiva: difatti se ad esempio su un paio di jeans presente su Zalando abbiamo 20 recensioni positive e 15 negative, la valenza oggettiva misurata sarà di +5, laddove invece la valenza soggettiva difficilmente avrà uguale valore, proprio per effetto della diversa percezione delle informazioni negative. D'altra parte, molto spesso i consumatori, in particolare quelli che sono alla ricerca del miglioramento della loro autostima, sono propensi a generare passaparola positivo riguardo ad esperienze di prodotti o servizi avute in prima persona, ma al contrario tendono a comunicare solo le esperienze negative avute da altre persone, fattore che di nuovo sottolinea il dualismo teorico tra passaparola positivo e negativo.

Riguardo a questo dualismo, Berger²⁰, riprendendo il secondo elemento del processo di comunicazione, ovvero il volume, afferma che i giudizi negativi non necessariamente si possono tradurre in fattori critici per le imprese, poiché anche la cattiva pubblicità aumenta la consapevolezza da parte dei consumatori del prodotto e del brand, attraverso la percezione di un volume più ampio di informazioni riguardo un particolare prodotto. La tipologia di sito riguarda invece il tema della fiducia e più nello specifico la principale distinzione del terzo elemento è tra siti di terze parti, come ad esempio Zalando per i prodotti nel settore del vestiario o Booking per i servizi legati alla ricezione alberghiera, e siti del brand del prodotto o servizio, come ad esempio il sito di Gucci o il sito di un Hotel. In particolare, risulta esserci molta più fiducia da parte dei consumatori per l'e-WOM di siti di terze parti per l'imparzialità che dovrebbe esserci nei giudizi presenti nelle recensioni. Infine l'elemento di esperienza nell'utilizzo del Web mostra un duplice effetto del passaparola elettronico sulla percezione e l'influenza del consumatore: da una parte possiamo affermare che la maggiore capacità di gestire questo strumento di comunicazione rende più consapevole la scelta e la selezione seguendo il criterio della valenza e soprattutto influenza in misura minore il comportamento del consumatore, laddove una scarsa capacità di muoversi nel Web potrebbe comportare un grande impatto dell'e-WOM sull'intensione di acquisto. D'altra parte, i consumatori con più esperienza

²⁰ The differential impact of brand loyalty on traditional and online word of mouth: The moderating roles of self-brand connection and the desire to help the brand, Jiska Eelen, Peren Özturan, Peeter W.J. Verlegh

digitale saranno influenzati in maniera più netta dalle fonti che ritengono credibili ed affidabili, proprio grazie alla capacità di discernere la qualità del comunicatore.

1.3 Wom, e-WoM e viral marketing

Nel panorama delle teorie che analizzano gli elementi di contatto e gli elementi disruptive tra WOM ed e-WOM, possiamo notare come ci siano pochi autori, come Filieri e McLeay²¹, che sostengono la possibilità di sovrapporre i due concetti, descrivendo l'e-WOM come una “*versione elettronica del passaparola tradizionale*”, non scorgendo invece la enorme opportunità per le imprese di aumentare esponenzialmente la creazione e la percezione di valore per e del consumatore, e di riflesso di aumentare le loro *revenues*. Dall'altra parte invece sono numerose le teorie che sottolineano le differenze sostanziali tra i due fenomeni, in particolare per quanto riguarda sette variabili: la credibilità, la privacy, la velocità di diffusione, l'accessibilità, la misurabilità, il tempo e la scalabilità. Secondo Cheung e Thadani²² la credibilità, intesa come risorsa di informazione è differente tra WOM ed e-WOM poiché l'anonimato del passaparola elettronico potrebbe essere visto come fonte di incertezza da parte del consumatore che ricerca l'informazione. Un altro fattore determinante nella differenza tra i due modelli è la Privacy, certamente maggiore nel passaparola tradizionale che in quello elettronico, laddove invece la velocità di diffusione risulta essere su due livelli completamente distinti tra WOM ed e-WOM, con quest'ultimo che ovviamente garantisce una velocità di diffusione spazio-temporale estremamente più elevata. L'accessibilità risulta essere maggiore nel passaparola elettronico, considerando ad esempio la possibilità da parte dei consumatori di poter consultare una recensione scritta in qualunque luogo e momento, vantaggio che è riassunto anche nella variabile tempo della analisi, notando oltre alla possibilità di consultazione asincrona delle informazioni dell'e-WOM, anche la persistenza del passaparola elettronico, ad esempio come una recensione su Amazon che ha potenzialmente una durata illimitata, laddove invece il passaparola tradizionale, essendo nella

²¹ A Literature Review of Word of Mouth and Electronic Word of Mouth: Implications for Consumer Behavior, Nuria Huete-Alcocer

²² The impact of electronic word-of-mouth communication: A literature analysis and integrative model, Christy M.K. Cheung, Dimple R. Thadani

maggior parte dei casi in forma orale, non permane per molto tempo. La misurabilità è la principale differenza tra le due forme di passaparola, poiché risulta essere determinante per la creazione di valore: difatti l'e-WOM, essendo completamente presente sul Web, consente alle imprese di collezionare le informazioni per poterle analizzare e di conseguenza creare una offerta customerizzata grazie agli algoritmi presenti nei *recommendation systems* che analizzeremo approfonditamente nel secondo capitolo. Collegandoci a questa ultima variabile analizzata, notiamo come il passaparola elettronico abbia una enorme scalabilità, potendo la sua analisi sistematicamente e con processi ingegnerizzati diventare una risorsa determinante per le imprese. Infine, una ulteriore distinzione tra le due tipologie di passaparola riguarda la scelta dell'informazione da comunicare: a pensarci bene la sincronicità e la spontaneità del WOM tradizionale, tipica di una conversazione tra conoscenti, si caratterizza come una comunicazione verbale che è espressione del pensiero del comunicatore. Nel caso del passaparola elettronico invece questo flusso, grazie alla asincronia del passaggio informativo, è filtrato dalla possibilità di revisione del messaggio e dal tempo a disposizione del comunicatore, il quale può così decidere cosa comunicare e come farlo. La Tabella 3 mostra le principali differenze appena descritte. Ai fini di rendere completa l'analisi sulle differenze risulta essere necessario sottolineare una ulteriore distinzione, stavolta tra l'e-WOM e il concetto di viral marketing, che vede una disparità di opinioni tra gli studiosi. Difatti se una corrente di pensiero identifica i due fenomeni come sovrapponibili, specificando che l'impegno attivo dell'impresa nel viral marketing si caratterizza esclusivamente nel creare un messaggio che sia quanto più possibile di facile comprensione e di contagiosa condivisione, d'altra parte vi sono esperti come Modzelewski e Wong²³ che affermano testuali parole: *"True viral marketing differs from word-of-mouth in that the value of the virus to the original consumer is directly related to the number of other users it attracts. That is, the originator of each branch of the virus has a unique and vested interest in recruiting people to the network."*

La differenza si sostanzia dunque, secondo la seconda corrente di pensiero, nelle caratteristiche del comunicatore, ed in particolar modo nel suo interesse a rendere una informazione virale, poiché se

²³ eWOM: Extant Research Review and Future Research Avenues, Anubhav Mishra and Satish S M

Tabella 3: Variabili in cui WOM ed e-WOM differiscono

	WoM	eWoM
Credibilità	Il destinatario dell'informazione conosce il mittente	Anonimato tra destinatario e mittente
Privacy	La conversazione è interpersonale e privata	L'informazione condivisa non è privata e può invece essere letta da chiunque e ovunque
Velocità di diffusione	Bassa velocità di diffusione dell'informazione	Alta velocità di diffusione dell'informazione
Accessibilità	Difficilmente accessibile	Facilmente accessibile
Misurabilità	Bassa se non assente	Alta attraverso sistemi come il machine learning e l'intelligenza artificiale
Tempo	Poco persistente e necessità di sincronicità	Persistente e asincrono
Scalabilità	Bassa se non assente	Molto alta
Scelta dell'informazione	Basso, con informazioni elargite spontaneamente in base all'esperienza in comunicazioni sincrone	Alto, con comunicatore che seleziona razionalmente cosa dire avendo tutto il tempo di organizzare le proprie idee

nel caso del viral marketing tale intenzione è ritenuta studiata a tavolino dalle imprese per avere un impatto positivo nella percezione dei consumatori, nel passaparola elettronico la diffusione delle informazioni è basata proprio sul disinteresse da parte del comunicatore nei confronti di un ritorno economico dell'impresa, fattore che aumenta la fiducia del ricevente nell'accogliere l'informazione come utile e veritiera.

1.4 L'importanza del tipo di linguaggio adottato dal comunicatore

Il tipo di linguaggio, analizzato attraverso la variabile della astrazione o concretezza dello stesso²⁴, assume un ruolo determinante nell'influenza che il passaparola ha sul consumatore, il quale a sua volta può avere un ruolo più o meno attivo nel processo di comunicazione a seconda della sua

²⁴ How language abstractness affects service referral persuasiveness Matteo De Angelis, Vito Tassiello, Cesare Amatulli, Michele Costabile

conoscenza dell'argomento. Per la nostra analisi risulta determinante categorizzare il tipo di linguaggio, attraverso il Linguistic Category Model, che prevede quattro principali livelli misurati, dal più astratto al più concreto: secondo questo modello difatti l'uso di aggettivi risulta essere il modo più generico e astratto nonché la forma di comunicazione che lascia maggiore libertà di interpretazione, laddove l'uso di verbi che esprimono un sentimento o uno stato psichico può essere considerato ad un livello di astrattezza meno elevato. Passando alle forme più concrete di linguaggio si hanno da un lato l'uso di verbi che esprimono azioni interpretabili, e ad un livello di concretezza maggiore i verbi che esprimono azioni fatte e concluse.

Da parte del comunicatore, secondo lo studio di Schellekens, l'uso di un linguaggio più astratto o più concreto è legato alla sua esperienza ed alla sua intenzione comunicativa: difatti se il consumatore ha avuto una esperienza positiva tenderà ad usare nel passaparola un linguaggio più astratto, nei termini espressi dal Linguistic Category Model, viceversa userà un linguaggio più concreto. Da parte del ricevente invece lo studio sottolinea che un linguaggio più astratto risulta essere più persuasivo e più efficace per destinatari che conoscono bene l'argomento di cui si parla, poiché quanto più è astratto il messaggio del comunicatore, tanto maggiore sarà la necessità che il ricevente abbia un ruolo attivo nella interpretazione del messaggio, e dunque la conoscenza della materia renderà più agevole l'elaborazione dell'informazione facendo leva sulla propria esperienza. Questo ruolo attivo sarà sviluppato grazie alla immaginazione mentale che scaturisce nel consumatore attraverso tre fattori, ovvero la vividezza, la quantità e l'elaborazione del messaggio.

CAPITOLO 2

2.1 Dal brick ad un click

Il Covid-19 ha accelerato in modo esponenziale il processo di digitalizzazione dei canali di marketing, e di pari passo ha reso il ruolo del passaparola elettronico sempre più centrale nel processo di acquisto, considerando che il 90% degli acquisti avviene in un negozio fisico ed il 60% delle decisioni dei consumatori su cosa e dove acquistare avviene online, dove è più semplice e rapido valutare i prodotti grazie alla grande mole di informazioni scambiate dai consumatori sulle piattaforme digitali. Possiamo dunque notare un fenomeno oramai comune che coinvolge le tipologie di fruizione dell'esperienza di acquisto, cioè il passaggio da canali *brick and mortar*, ovvero i classici store fisici dislocati tra le strade delle città, a *brick and click*, ossia una forma ibrida che comprende la possibilità di acquistare i prodotti in negozi fisici o virtuali, da molti ritenuta la giusta via di mezzo considerando da un lato la accelerazione digitale che integra la tecnologia allo retailer fisico, ma dall'altro la propensione di molti consumatori ad acquistare negli store fisici, dove tra questi il 65% per provare il prodotto, il 52% per il fattore gratificazione ed il 33% per sicurezza, fino ai canali pure-click, esclusivamente presenti sulle piattaforme digitali. In questo contesto gli smartphone, secondo Netcom²⁵ utilizzati per effettuare il 51% degli acquisti online, e tutti gli strumenti digitali che supportano quotidianamente la curiosità, le scelte e i dubbi degli individui, hanno avuto un ruolo centrale nel corso dell'ultimo anno, cercando di garantire la *serendipity degli acquisti e dei consumi*²⁶ che risulta essere la base imprescindibile dell'emotività dei consumatori durante il processo di acquisto.

Come accennato, questa rivoluzione digitale dell'esperienza di acquisto ha un impatto determinante anche sull'esperienza post acquisto, e dunque sulla modalità dei consumatori di comunicare le proprie sensazioni e i propri giudizi agli altri individui che sono in una fase precedente all'acquisizione del prodotto, e che dunque supportano in misura sempre maggiore le proprie scelte grazie alle informazioni generate dall'e-WOM. Difatti la ricerca online è diventata il primo stadio del processo di acquisto, che può essere diviso, come mostra la Figura 2, in tre fasi considerando la variabile della verità emozionale del consumatore²⁷: la prima fase, dopo lo stimolo iniziale, è proprio la Zero moment

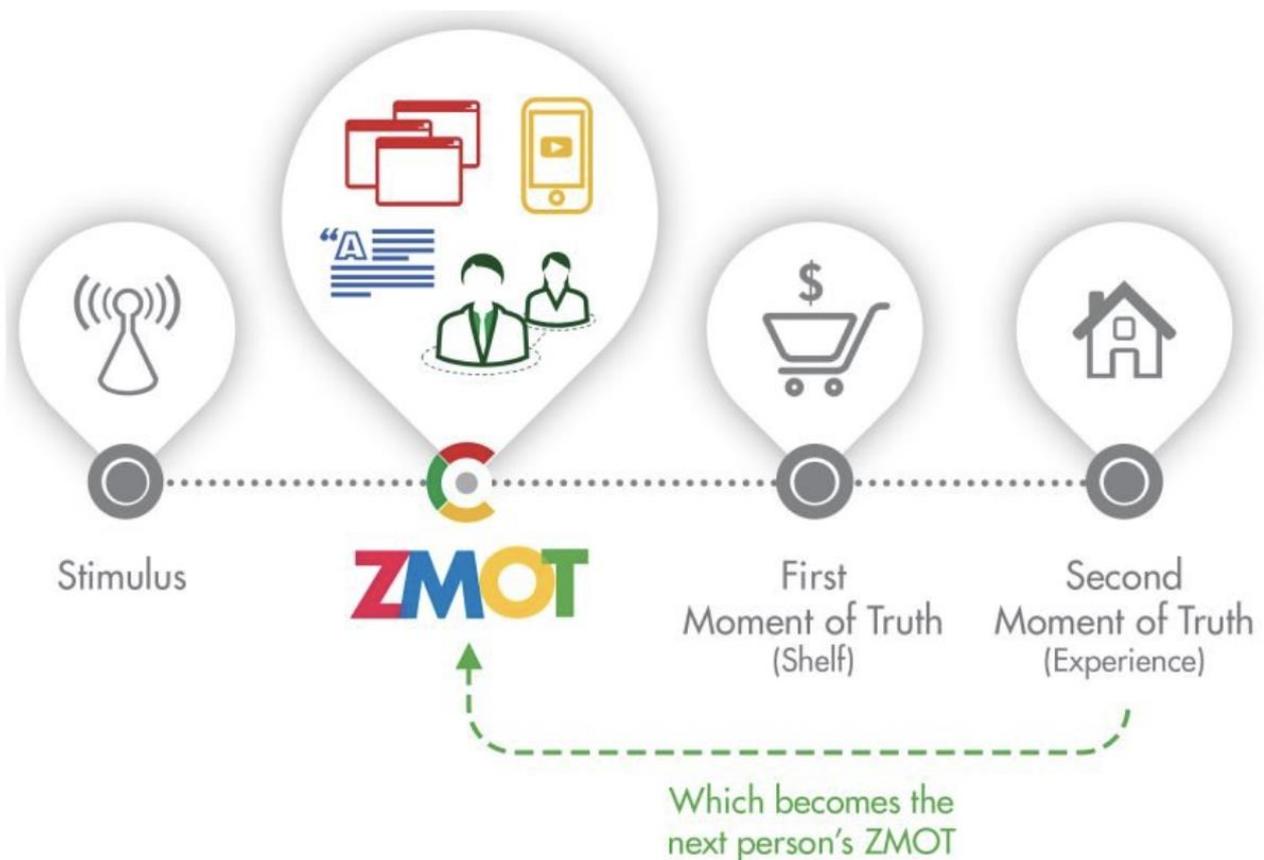
²⁵ <https://www.freeway.it/analytics/marketing-operativo-data-driven>

²⁶ Il Sole 24 Ore, 25 Marzo 2021, Un nuovo modo di consumare in bilico tra negozi fisici e virtuali, Costabile e Portu

²⁷ <https://enricomtomassi.com/zmot-social-conversione/>

of trust, un concetto elaborato da Google nel 2011, nel quale l'individuo attraverso le informazioni generate dall'e-WOM compie la decisione di acquisto; la seconda fase, la First moment of trust è invece il momento in cui il soggetto compra fisicamente il prodotto ed infine la Second moment of trust è la fase nella quale il consumatore fa esperienza del prodotto, e che poi attraverso la stesura del suo giudizio sulle piattaforme digitali, trasforma la terza fase nella prima degli altri consumatori.

Figura 2: Fasi processo decisionale



EMT 2019

Questo fenomeno comporta da parte delle imprese la necessità di avere un elevato grado di attenzione nella gestione di questo mezzo di comunicazione indiretto, con l'obiettivo di renderlo il più diretto possibile. Rowley²⁸ ha affermato a riguardo che le imprese dovrebbero implementare blog e community brandizzate, facendo leva sulla loyalty dei propri consumatori, e avendo una gestione quanto più possibile trasparente, soprattutto riguardo le review negative. La gestione dell'e-WOM da

²⁸ The Effect of E-WOM on Customer Purchase Intention, Noraini Sa'ait, Agnes Kanyan and Mohamad Fitri Nazrin

parte delle imprese risulta essere un punto cruciale del rapporto con i consumatori a causa della maggiore diffidenza di questi ultimi nell'acquisto, come conseguenza di congiunture storiche come la "Grande recessione" degli anni 2007-2008 o come la attuale crisi che ha portato ad una drastica riduzione dei consumi, contesto che sta dando sempre maggiore importanza all'e-WOM, come sottolineato dal dato che vede il 70% degli italiani consultare il Web per avere informazioni sul prodotto o servizio che gli interessa. Ci sono numerosi casi di mancanza di trasparenza nella gestione delle recensioni online, come ad esempio dimostra il fenomeno dei 3 domini al giorno con la parola Coronavirus o Covid²⁹ con recensioni false sulla vendita di prodotti che riescono a proteggere o cosa ancora più grave a curare i pazienti affetti dal virus, oppure come il caso di B6, l'app lanciata dalla banca scozzese RBS, che prima ancora del lancio ha riscontrato sul sito della banca numerose recensioni, provocando una perdita di fiducia irrecuperabile da parte dei consumatori. Ed è proprio per casi come questi che la Competition and Market Authority lavora per garantire un alto livello di trasparenza, attraverso l'implementazione di sistemi come Trustpilot che migliorano la trasparenza informativa delle recensioni con un attento controllo della gestione delle recensioni da parte delle imprese, considerando che il 47% dei consumatori non si fida dei marchi che propongono solo recensioni positive poiché il 40% di loro crede che queste aziende manipolino le recensioni, cercando dunque di annullare comportamenti scorretti che portano ad una asimmetria tra benefici per le imprese nel breve periodo e perdita di fiducia da parte dei consumatori nel medio lungo periodo. Inoltre, fare una scrematura dei giudizi lasciati dai clienti sui siti online delle aziende risulta essere controproducente anche perché mantenere le recensioni negative, oltre che garantire la fiducia nella fonte, equilibra anche le aspettative dei consumatori prima dell'acquisto, componente che se troppo elevata rischia di diminuire il tasso di soddisfazione della clientela.

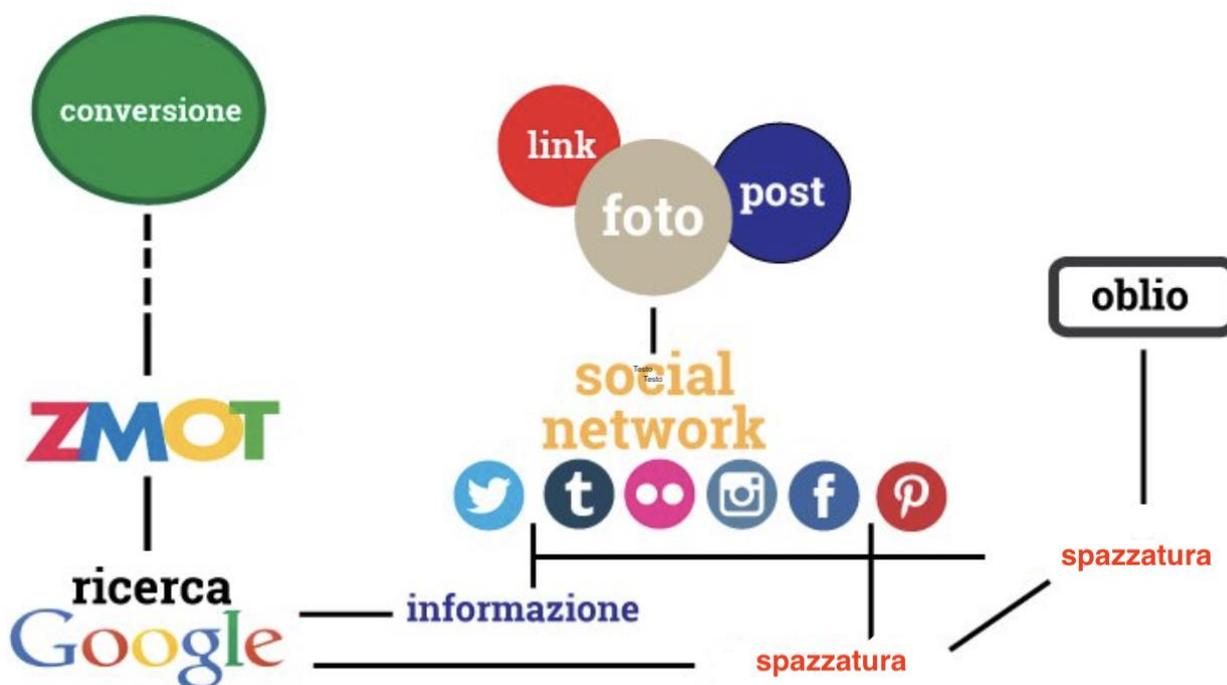
2.2 La tecnologia a supporto dell'impresa

Avendo analizzato in che modo il virtuale sta affiancando o in alcune circostanze sostituendo il mondo reale, e avendo ben chiaro il dualismo WOM e-WOM, la vera sfida per le imprese risulta capire come implementare sistemi che elaborino attivamente le sensazioni dei consumatori, generando una rete di informazioni sotto forma di suggerimenti, che indirizzino le scelte degli individui durante i processi decisionali di acquisto, come fosse una forma di passaparola elettronico, che nasce però dalla crescente capacità delle imprese di incamerare ed analizzare i dati, internalizzata

²⁹ <https://it.businessinsider.com/ascoltate-sempre-i-vostri-clienti-la-fiducia-dei-consumatori-si-basa-su-recensioni-aperte-e-trasparenti/>

nelle strategie di comunicazione e di canale delle aziende. Un metodo intermedio tra l'e-WOM dei consumatori ed i sistemi di raccomandazione delle imprese potrebbe essere rappresentato dai social network, che sono oggi una grande opportunità per sostenere il tasso di conversione dei brand e la fidelizzazione dei soggetti che hanno già avuto esperienza del prodotto, e di conseguenza per sostenere la creazione di passaparola positivo. Il rapporto che le imprese riescono a creare con i propri *follower* sui social è dovuto alla capacità empatica ed inclusiva del canale, che risulta essere dunque su un piano completamente differente rispetto ai meccanismi automatizzati, che analizzeremo nei prossimi paragrafi, i quali grazie ad elaborati algoritmi, all'AI ed al Machine Learning, personalizzano l'offerta ed il layout dei siti in base ai dati acquisiti sul consumatore. Come mostra la Figura 3 gli individui che entrano a contatto con un post, una storia o un *tweet* del profilo social di una impresa possono avere due prime tipologie di impressioni: da un lato lo stimolo può essere percepito come una informazione utile, mentre dall'altro come spazzatura, e se l'informazione o la spazzatura generano ulteriore spazzatura dopo una analisi dell'utente, il contenuto è destinato a finire nell'oblio; dall'altra parte se invece il contenuto desta la curiosità dell'individuo, questa informazione verrà approfondita fino ad entrare in contatto con la Zero moment of trust descritta nel paragrafo precedente, dal quale può nascere l'interesse nel provare un prodotto del marchio.

Figura 3: Percezione dei contatti con social network

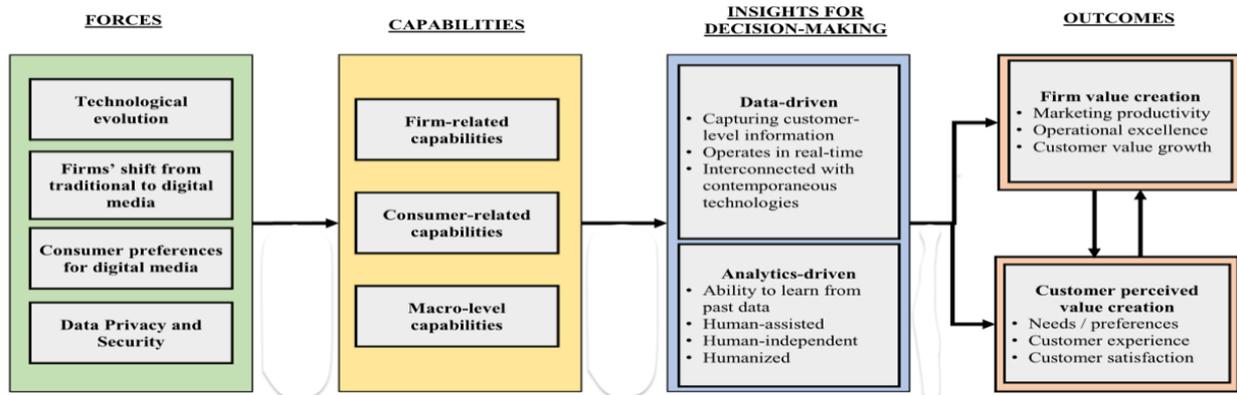


La *digital analytics*, intesa come la capacità delle imprese di utilizzare i dati acquisiti dal comportamento del consumatore ed analizzarli per ricavarne valore, è oramai diventata la discriminante per lo sviluppo delle aziende in ogni tipologia di settore industriale, come ad esempio quello delle banche. Difatti il caso di ZestFinance³⁰ risulta essere rappresentativo per capire quanto i dati oggi rappresentino un valore economico: questo è un algoritmo che incrocia i dati raccolti tramite i cookies, basandosi ad esempio sul tempo che l'individuo ha passato a leggere i termini del contratto, con altre caratteristiche del richiedente del prestito per dare un rating del rischio. La centralità delle nuove tecnologie legate all'analisi dei dati, mostrata anche dalle affermazioni dei Chief Marketing Officers interpellati da una analisi di IBM, i quali hanno affermato di essere maggiormente concentrati sulla esperienza del consumatore piuttosto che sul prodotto, sta accelerando anche il processo di personalizzazione dell'offerta, oramai considerato real time rispetto all'immagazzinamento e alla analisi dei dati. La teoria di Gupta, Leszkiewicz, Kumar, Bijmolt e Potapov³¹ crea a riguardo una base concettuale utile a concepire l'importanza delle nuove tecnologie per creare gli insights fondamentali alle imprese per generare offerte personalizzate, che risultino agli occhi dei consumatori una forma di comunicazione il più naturale possibile, quasi come fosse un passaparola tra conoscenti, in cui ora il ruolo del comunicatore è assunto dall'impresa e non più dagli altri consumatori attraverso AI, Machine Learning e Big Data, i quali supportano la creazione degli algoritmi alla base dei sistemi di raccomandazione. In questo framework teorico, riassunto nella Figura 4, ci sono quattro forze che sostengono l'integrazione della tecnologia nelle strategie delle imprese, ovvero l'evoluzione tecnologica, i digital media, le preferenze dei consumatori nell'usare i digital media ed infine la privacy. Le tre capacità che fanno la differenza nell'implementare la tecnologia sono le capacità delle imprese, dei consumatori e le capacità a livello macro, le quali portano all'analisi di *insights* attraverso i quali si genera valore sia per le imprese, in termini di prodotti più aderenti alle necessità dei clienti, e quindi ad un aumento dei ricavi, e di conseguenza un aumento della percezione del valore da parte dei consumatori.

³⁰ Digital Analytics: Modeling for Insights and New Methods Shaphali Gupta & Agata Leszkiewicz & V. Kumar & Tammo Bijmolt & Dmitriy Potapov

³¹ Digital Analytics: Modeling for Insights and New Methods Shaphali Gupta & Agata Leszkiewicz & V. Kumar & Tammo Bijmolt & Dmitriy Potapov

Figura 4: Framework teorico per integrare la tecnologia nell'impresa



Digital Analytics: Modeling for Insights and New Methods Shaphali Gupta & Agata Leszkiewicz & V. Kumar & Tammo Bijmolt & Dmitriy Potapov 2020

In questo contesto il tema della privacy è uno dei più discussi e di difficile trattazione, poiché ci si muove in un labile confine tra l'efficienza dei sistemi di raccomandazione e quindi offerte più aderenti ai bisogni dei consumatori, e la sensibilità dei dati privati. Shoshanna Zuboff³² nel suo libro intitolato *“Il capitalismo della sorveglianza”* sostiene la tesi secondo cui i colossi del Web come Google e Facebook siano in continua lotta con i sistemi di protezione della privacy online, e con le annesse legislazioni, poiché ritenute un grande ostacolo alla raccolta dei dati dei loro utenti, e alla loro possibile vendita alle imprese. Secondo Gupta, Leszkiewicz, Kumar, Bijmolt e Potapov³³ la rivoluzione digitale ha introdotto tre nuove problematiche riguardanti la gestione dei dati degli individui: il primo riguarda la capacità delle imprese di sapere in anticipo quale sarà il comportamento di acquisto del consumatore, potendo in questo modo ad esempio applicare discriminazioni di prezzo del terzo tipo, ovvero con prezzi diversi per ogni individuo in base alla sua disponibilità a pagare; il secondo invece riguarda la impossibilità da parte delle imprese di internalizzare gli eventuali danni arrecati ai consumatori per un attacco alla propria banca dati, a causa delle difficoltà nel tracciamento degli hacker; infine il terzo problema riguarda la tendenza delle imprese a concordare con i consumatori policy di protezione dei dati che poi non riescono a mantenere, poiché risulta molto complicato per i consumatori verificare la corretta gestione delle informazioni personali. L'importanza della questione viene sottolineata anche dai numerosi casi di incorretta protezione dei

³² Shoshanna Zuboff, *Il capitalismo della sorveglianza*

³³ Digital Analytics: Modeling for Insights and New Methods Shaphali Gupta & Agata Leszkiewicz & V. Kumar & Tammo Bijmolt & Dmitriy Potapov

dati personali detenuti dalle imprese, come mostra il dato secondo il quale nel 2016 negli Stati Uniti ci sono stati oltre 1000 casi di violazione dei diritti sulla privacy, ben il 40% in più rispetto all'anno prima.

2.2.1 Big data

Nei paragrafi precedenti ci siamo serviti numerose volte del concetto di big data, intesi come l'ingente mole di dati che vengono raccolti e analizzati per implementare forme di personalizzazione di massa nel contesto dell'industria 4.0, attraverso nuove tecnologie come l'*Internet of Things*, ovvero le interconnessioni tra oggetti in rete, pensiamo ad esempio alla innumerevole quantità di sensori che ogni istante misurano le nostre attività, e grazie alla intelligenza artificiale e al Machine Learning, ovvero la capacità delle macchine di apprendere autonomamente durante le loro attività. La raccolta e l'analisi di questi dati diventa determinante per lo sviluppo da parte delle imprese di sistemi di raccomandazione che risultano oggi alla base degli algoritmi della maggior parte delle aziende presenti in rete. L'analisi delle imprese del comportamento del consumatore, e pensiamo ad esempio alle reazioni agli stimoli come e-mail o newsletter o alle modalità e ai tempi di acquisto negli store fisici, come ad esempio la *basket analysis*, e online, ha lo scopo di far comunicare le persone inconsapevolmente all'impresa le loro preferenze, verso le quali questa rivolge l'attenzione, in termini ad esempio di sviluppo di prodotto. Questa analisi dei dati viene condotta attraverso KPI di copertura come quanti utenti hanno ricevuto il messaggio dell'impresa e quale impressione hanno avuto, che si riferiscono alle fasi iniziali del marketing funnel, e con i KPI di interazione, che invece fanno riferimento alle fasi finali del funnel, come l'engagement e l'acquisto del prodotto. Difatti come affermato da Vittorio Cretella³⁴, Ceo di P&G, imprese come la sua stanno diventando aziende di dati, e questo per cercare di anticipare il mercato, suggerendo attraverso i sistemi di raccomandazione i prodotti ai consumatori, proponendo così una alternativa all'e-WOM, che però risulta essere sotto completo controllo delle imprese, a differenza del passaparola, cercando di rendere questi suggerimenti i più naturali possibile, guadagnando in tal modo la fiducia dei consumatori.

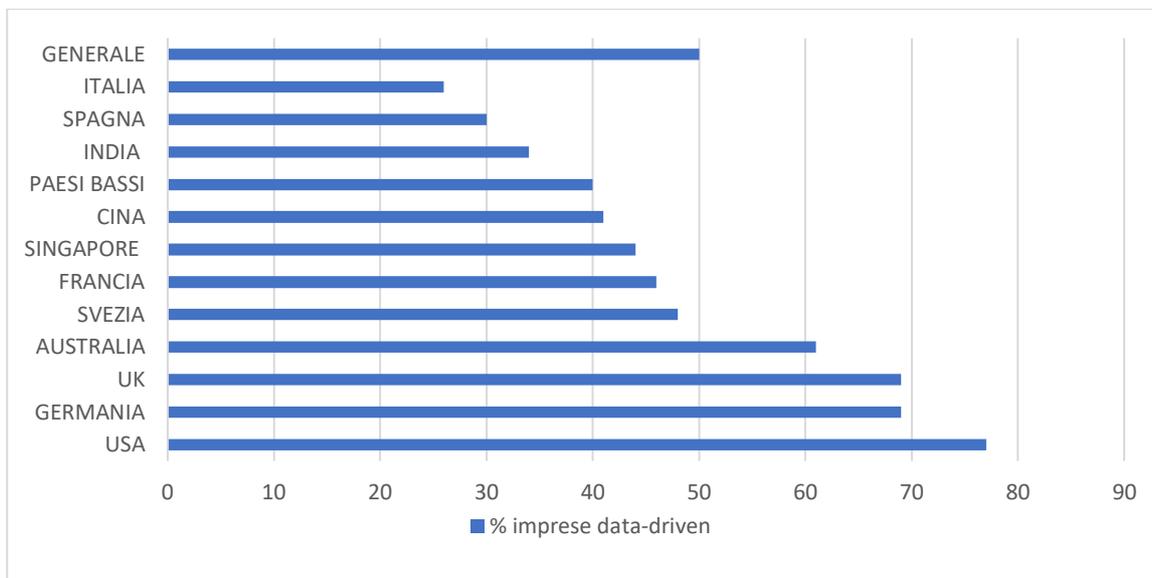
Secondo uno studio condotto dal Sole 24 Ore³⁵, il mercato europeo dei dati ha avuto una crescita esponenziale negli ultimi anni, ed in particolare se nel 2015 sei milioni di persone lavoravano con i dati, nel 2016 questo numero è aumentato a 6.160.000, arrivando a 10.430.000 nel 2020, con un tasso di crescita annuale del 14,1%, con 255.000 aziende coinvolte nel 2016, arrivando a più di 350.000

³⁴ Il Sole 24 Ore, Big data, noi siamo quello che compriamo. E il supermercato lo sa meglio di noi, 12 Febbraio 2021, Luca Tremolada

³⁵ <https://www.infodata.ilsole24ore.com/2017/05/04/quanto-vale-mercato-europeo-dei-dati/>

nel 2020. Se nel 2016 il mercato europeo dei dati valeva 300 miliardi, questo valore è più che raddoppiato nel 2020, superando i 700 miliardi. Inoltre, in Italia solo il 26% delle aziende sfrutta i dati che ha a disposizione e quelle che lo fanno aumentano i loro profitti del 22%, mentre negli altri paesi, come è possibile notare dalla Tabella 4, si hanno valori più elevati, con la media mondiale che si attesta su una impresa su due considerata *data-driven*.

Tabella 4: Imprese data driven in Europa



Il Sole 24 Ore 2021

Questi numeri ci fanno capire perché i Big Data sono definiti l'oro nero, o il petrolio del nuovo millennio, ma è necessario sottolineare che questi sono dati grezzi, per cui per ottenere davvero grandi ricavi economici e reputazionali, le imprese devono implementare sistemi che li analizzino e che traggano informazioni utili. Erevelles, Fukawa e Swayne³⁶ hanno creato un framework concettuale basato sulla centralità delle risorse per sfruttare al meglio i dati a disposizione: per ottenere importanti vantaggi economici, concorrenziali e di reputazione con i dati, è necessario, secondo questa teoria, possedere risorse preziose, rare, imperfettamente imitabili e sfruttabili dall'impresa. Le risorse poi possono essere classificate come fisiche, umane e organizzative, dove le risorse fisiche comprendono tutte i software e gli hardware necessari per lo storage e la elaborazione dei dati, le risorse umane sono le persone con competenze specifiche nella gestione ed analisi dei dati, mentre le risorse organizzative riguardano la struttura dell'impresa, ed in particolar modo in quale maniera i *data*

³⁶ Big Data consumer analytics and the transformation of marketing Sunil Erevelles, Nobuyuki Fukawa, Linda Swayne

scientists sono inseriti all'interno dell'organigramma e del flusso decisionale. Infine, per comprendere a fondo il concetto di Big Data è necessario esporre la teoria delle 3 V, considerate le caratteristiche che descrivono al meglio questo fenomeno: queste sono la velocità, il volume e la varietà. La velocità consiste nel timing sempre più accelerato di creazione di dati, per il volume invece si stima che ogni due anni la necessità di hardware per lo storage dei dati raddoppia, considerando che una grande accelerazione è stata data dall'IoT, ovvero dalla connessione in rete di oggetti aventi sensori, con 32 miliardi di oggetti connessi in rete nel 2020; infine la varietà è una delle caratteristiche che maggiormente differenzia i Big Data dai dati strutturati, in quanto i primi avendo una enorme varietà di fonti sono nella maggior parte dei casi destrutturati o parzialmente strutturati. Ma queste tre caratteristiche ci portano ad aggiungere altre due V alla analisi, ovvero la veridicità e il valore dei dati, proprio perché il crescente volume, la sempre maggiore velocità e varietà di dati rendono necessario una attenta scrematura degli stessi.

2.2.2 AI e Machine Learning

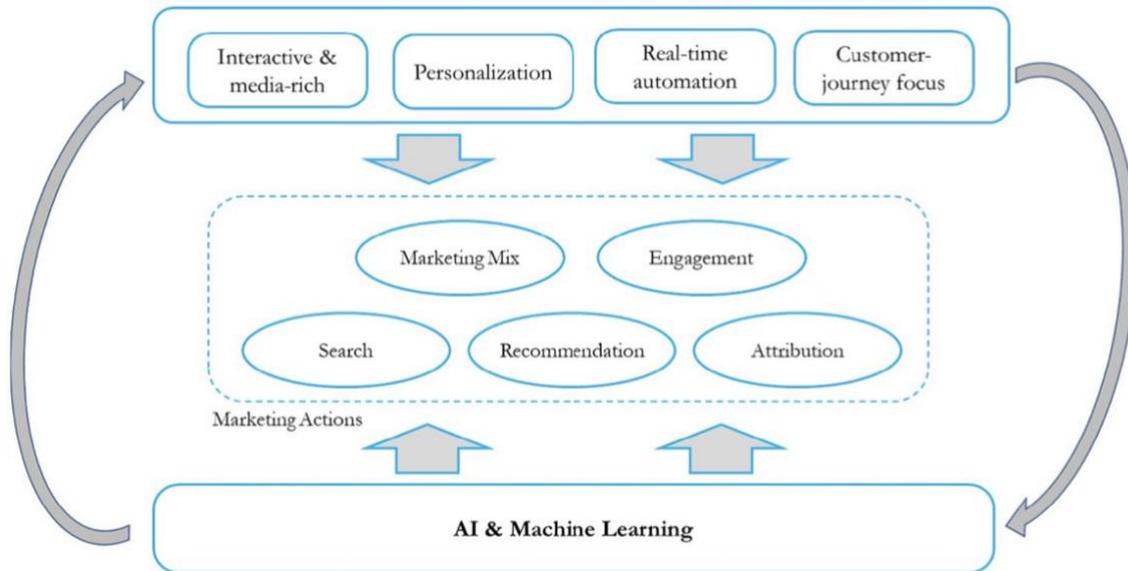
Prima di analizzare nello specifico nel prossimo paragrafo i sistemi di raccomandazione, proposti come lo strumento implementato dalle imprese per contrastare la impossibilità di controllo delle stesse sul passaparola, elettronico e non, cercando allo stesso tempo di mantenere la percezione di disinteresse del comunicatore, il consumatore nell'e-WOM e l'impresa nei *recommendation systems*, nel fornire consigli nel processo di acquisto, è necessario svolgere una analisi delle tecnologie che rendono possibile l'elaborazione dei dati, ovvero l'intelligenza artificiale e il Machine Learning. Pensiamo ad un motociclista che avvia una ricerca per acquistare un nuovo casco. Il primo passo sarà compiere una ricerca su internet per orientarsi sul prezzo, sui modelli e sui brand, farà una scrematura attraverso i filtri di siti di e-commerce, poi dopo aver individuato un prodotto che lo soddisfa come linea e caratteristiche, e che rientra nel suo range di prezzo lo andrà a provare in uno store fisico, per capire la taglia giusta, fenomeno definito ROPO, ovvero leggibile come *research online and purchase off-line*, o viceversa come *research off-line and purchase online*, e poi magari tornerà su internet per cercare lo stesso prodotto ad un prezzo scontato rispetto al negozio, compiendo l'acquisto dopo aver sentito anche il parere di un gruppo di *riders* suoi amici. Nel corso del processo di acquisto di questo motociclista attraverso le sue ricerche ha lasciato diverse impronte sulle sue preferenze e sul suo comportamento, generando per un solo acquisto una grande quantità di dati, gli stessi che sono il motore che attiva gli algoritmi di intelligenza artificiale. L'AI è stata definita nel corso degli anni da

grandi scienziati³⁷ come Alan Turing, che nel 1950 si chiese se le macchine potessero pensare, o come McCarthy, il quale affermò che l'AI è *“That of making a machine behave in ways that would be called intelligent if a human were so behaving”*. Il Machine Learning invece, definito da Mitchell come *“A computer program that is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience”*, è considerato essere una applicazione specifica dell'AI, e una alternativa o una integrazione nel marketing alle analisi econometriche, le quali non risultano esaustive con una grande mole di dati. A riguardo il framework concettuale proposto dalla teoria di Ma e Sun³⁸, valuta l'AI e il Machine Learning considerando l'intero *customer journey*, e ponendo l'accento sul metodo di analisi, sui dati che devono essere ricchi e ottenuti da fonti differenti, su problemi rilevanti e su una integrazione con le principali teorie di Marketing. La teoria propone diverse classificazioni del Machine Learning, il quale può essere supervisionato, dove sia la variabile dell'input, ovvero il set di dati da cui si parte, che l'output sono rilevati dagli algoritmi, o non supervisionato, dove al contrario solo l'input è conosciuto, semi supervisionato, caratterizzato da una conoscenza completa dell'input, e una parziale degli output, e attivo, nel quale la scarsa quantità di dati iniziali comporta una sfida di efficienza per gli algoritmi che devono ottenere la massima potenza predittiva sul comportamento del consumatore con poche risorse, situazione tipica quando si studia il comportamento di un nuovo cliente del brand. La teoria analizza inoltre l'AI, sottolineando come questa abbia introdotto numerose tendenze nelle imprese, quali l'automazione, la personalizzazione di massa, un rapporto interattivo e ricco di contenuti multimediali e la focalizzazione sul *customer journey*, che hanno portato al ciclo positivo mostrato in Figura 5. Gli innumerevoli scambi multimediali che ci sono tra imprese e consumatori attraverso i social media portano ad avere la necessità di implementare algoritmi che massimizzino l'engagement e la fidelizzazione da questi canali, sfruttando le diverse forme in cui questi contatti avvengono, come video, audio, immagini o testo. La personalizzazione di massa sfrutta la possibilità di avere informazioni precise su ogni singolo individuo ottenendo una targetizzazione che deve essere non solo corretta in termini qualitativi, ma che deve essere effettuata nel momento e nel contesto giusto.

³⁷ Machine learning and AI in marketing – Connecting computing power to human insights, Liye Ma, Baohong Sun

³⁸ Machine learning and AI in marketing – Connecting computing power to human insights, Liye Ma, Baohong Sun

Figura 5: AI nelle imprese



Machine learning and AI in marketing – Connecting computing power to human insights, Liye Ma, Baohong Sun 2020

Da questo ultimo punto si collega il tema della automazione real time, che ottimizza la rapidità di raccolta ed elaborazione di grandi quantità di dati che solo gli algoritmi hanno, per implementare sistemi di raccomandazione che sfruttano al meglio la finestra temporale in cui il consumatore ricerca un prodotto nell'e-commerce o in uno store fisico attraverso sistemi che utilizzano i cookies o la geolocalizzazione. Tutto questo è possibile solo se l'attenzione delle imprese si focalizza sull'intero customer journey, attraverso l'analisi del comportamento del consumatore in tutte le sue fasi. Questi trend legati all'AI e al Machine Learning devono essere integrati con le principali azioni di marketing quali la ricerca, il marketing mix, l'attribuzione, l'engagement ed infine la raccomandazione. Difatti i sistemi di intelligenza artificiale oramai controllano le variabili del marketing mix, come il prezzo che sempre più spesso viene definito usando analisi quantitative basate sui dati, o il customer engagement che come descritto è oramai un obiettivo dei team di Marketing delle imprese raggiungibile solo con il supporto dei dati. Stesso ragionamento vale per la ricerca di marketing e per l'attribuzione, considerati gli innumerevoli punti di contatto presenti oggi tra impresa e consumatori grazie ai social media.

2.3 Recommendation systems

Gli shop online non sono più uno spreco di tempo bensì favoriscono il benessere dei consumatori in termini economici e di tempo investito nella ricerca di informazioni e nell'acquisto di un prodotto. Le imprese, dunque, per battere sul tempo la concorrenza hanno la necessità di implementare algoritmi che studino il profilo di ogni utente per ottenere un'offerta personalizzata, che ricalchi i bisogni del consumatore in uno specifico momento. Tale obiettivo, ottenuto attraverso i sistemi di raccomandazione, è inoltre come anticipato nei precedenti paragrafi uno strumento utile alle imprese nell'indirizzare le scelte del consumatore, così come accade nel passaparola, con la differenza che se le recensioni o i blog offrono pareri disinteressati rispetto al prodotto e al marchio oggetto della discussione, i suggerimenti dei siti di e-commerce di brand specifici, come il sito di Gucci, o generalisti, come invece può essere Amazon, possono essere percepiti dal consumatore come una forma di passaparola che sottende interessi specifici dell'azienda. Possiamo dividere i sistemi di raccomandazione in due macro-categorie: *content based* e *collaborative filtering*, dove la prima tipologia basa il proprio algoritmo sulle caratteristiche dei prodotti scelti dai consumatori, e la principale criticità è data dalla quantità di caratteristiche necessarie affinché il sistema lavori correttamente. Il *collaborative filtering* invece raggruppa attraverso una *cluster analysis* i consumatori in base alle loro scelte, e le due principali criticità riguardano la necessità di avere pareri espliciti degli individui sui prodotti per avere un livello di accuratezza maggiore, e la scalabilità, poiché come il primo sistema anche questo lavora meglio ed in modo più veloce con una quantità di dati più ampia.

L'analisi svolta da Hung³⁹ riguardo i *recommendation systems* ha l'obiettivo di costruire un sistema basato sulla tassonomia dei prodotti e sulla classificazione dei consumatori. Tenendo conto del principio di Pareto per cui il 20% dei clienti di una impresa genera l'80% dei suoi profitti, questo sistema di raccomandazione ha l'obiettivo di aumentare il tasso di ritenzione dei consumatori visibili online, ovvero quei consumatori di cui hanno molte informazioni e che quindi permettono una targettizzazione *one to one*, in modo da soddisfare al meglio i bisogni dell'individuo ed aumentarne la percezione di valore del brand. Le caratteristiche analizzate da questa teoria per individuare i target di consumatori riguardano la variabile comportamentale, ed in particolare la *brand loyalty* e il tasso

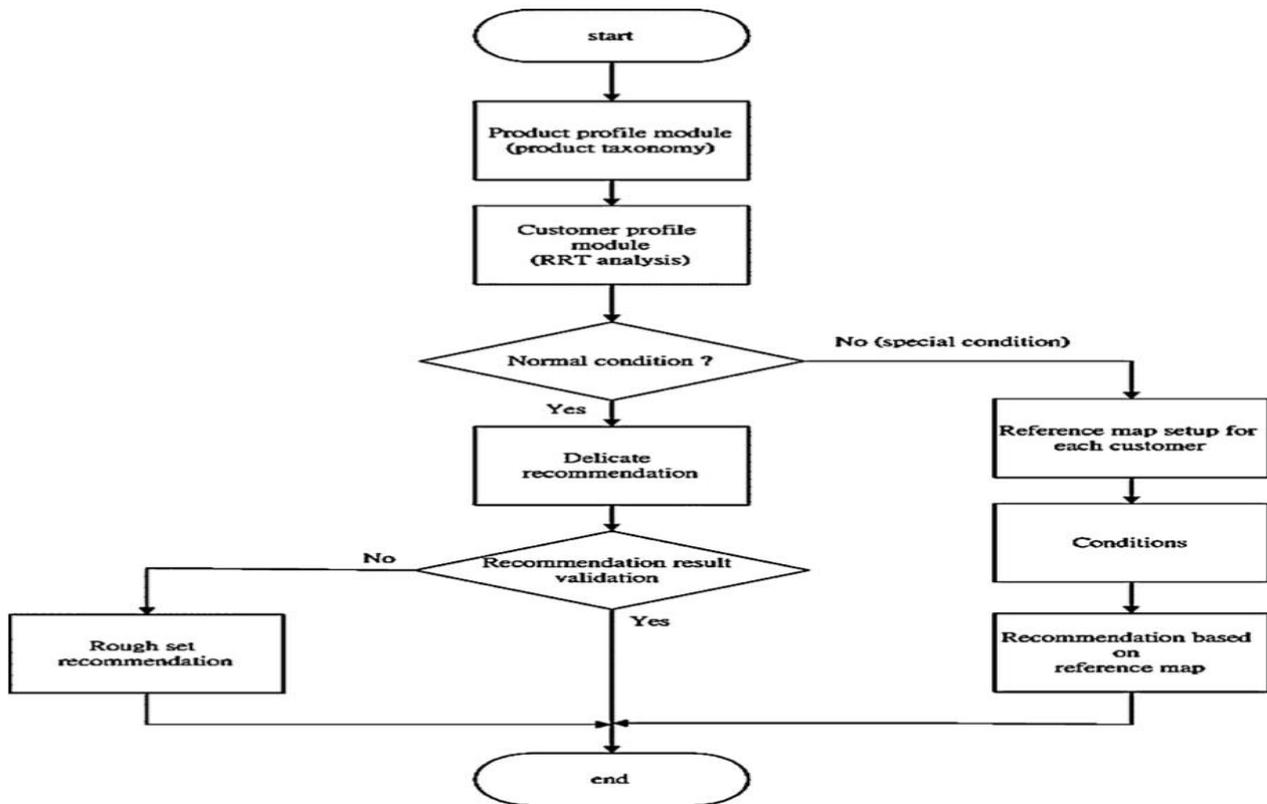
³⁹ A personalized recommendation system based on product taxonomy for one-to-one marketing online, Lun-ping Hung

di acquisto. Questo sistema di raccomandazione si basa, come la maggior parte di questi algoritmi, sullo studio delle preferenze dello specifico consumatore, incrociate a quelle di consumatori con gusti simili. In particolar modo la Flow-chart del sistema di raccomandazione proposta dalla teoria e mostrata nella Figura 6, si compone di diversi step che portano alla creazione di un algoritmo che cambia le sue caratteristiche a seconda della condizione dell'ambiente. Il primo step, si occupa di creare un modulo del profilo del prodotto, analizzando la tassonomia dei prodotti venduti dal brand, che è identica per qualsiasi individuo, ma varia a seconda del venditore. Lo step successivo consiste nell'analisi della tendenza al riacquisto recente di ogni consumatore, la così detta *RRT analysis*, ed è a questo livello che il sistema di raccomandazione caratterizza diversamente il profilo di ogni utente. Difatti in questo step si definiscono il livello di *brand loyalty* e il tasso o frequenza di acquisto, utilizzando una formula che sintetizza queste due caratteristiche del comportamento del consumatore, laddove un risultato più alto indica acquisti più recenti e frequenti. Infine, se la condizione è normale, intesa come una situazione in cui sono disponibili grandi quantità di dati su un consumatore che quindi ha un alto valore di RRT, viene attivata una raccomandazione delicata, la quale se non produce risultati sufficienti, ovvero se la raccomandazione avviene per un nuovo cliente o per un nuovo prodotto, viene implementata con una raccomandazione grezza. Se la prima tipologia si basa su una attenta analisi degli acquisti passati dell'individuo, la seconda invece ha un livello di incertezza maggiore, in quanto raccomanda anche tipologie di prodotti non acquistati recentemente dal consumatore, potendo calcolare le sue preferenze su un set scarso o assente di dati. Abbiamo definito i sistemi di raccomandazione come una forma di passaparola internalizzato dall'impresa, che suggerisce i prodotti ai clienti in base ad una analisi delle loro preferenze. Va sottolineato che un sistema di raccomandazione funziona se vengono sfruttati il maggior numero di *touchpoints* digitali con i consumatori, e un fallimento dei *recommendation systems* provocherà ovviamente una catena negativa di passaparola tradizionale ed elettronico. Un secondo modello di raccomandazione che è descritto per avere un esempio di come è generato un *recommendation system*, è quello di Behera, Gunasekaran, Gupta, Kamboj, Bala⁴⁰, che si basa su due punti chiave: il design dell'e-marketplace, il quale comprende la classificazione dei clienti, dei prodotti, della proposta unica di vendita per ogni individuo, e lo sviluppo di un sistema di raccomandazione dei prodotti considerando le diverse

⁴⁰ Personalized digital marketing recommender engine Rajat Kumar Behera, Angappa Gunasekaran, Shivam Gupta, Shampy Kamboj, Pradip Kumar Bala

categorie di individui, attraverso tecniche di vendita come il *crossselling*, *l'up-selling* o il *need-satisfaction selling*.

Figura 6: Schema sistema di raccomandazione proposto da Lun-ping Hung



A personalized recommendation system based on product taxonomy for one-to-one marketing online, Lun-ping Hung

Le variabili che costituiscono la raccomandazione personalizzata sono classificate in variabili di trattamento e variabili controllate, dove tra le prime sono presenti i consumatori, i prodotti e la strategia di vendita, mentre tra le seconde sono composte dalla proposta unica di vendita della proposta di Marketing dell'impresa. La teoria parte da tre ipotesi, partendo dal presupposto che implementando il sistema si avrà un livello maggiore di raccomandazione personalizzata digitale, che quindi comporterà:

1. Un livello più alto della media del valore di ogni ordine
2. Revenues più alte
3. Un numero più alto di prodotti per ordine.

Il framework di questo sistema di raccomandazione richiede la clusterizzazione di consumatori, Figura 7, prodotti, Figura 8 e proposte uniche di vendita, Figura 9, per aumentare il tasso di conversione degli utenti in clienti fedeli.

Figura 7 (sinistra), 8 (destra) e 9 (in basso): Clusterizzazione dei consumatori, dei prodotti e proposte uniche di vendita

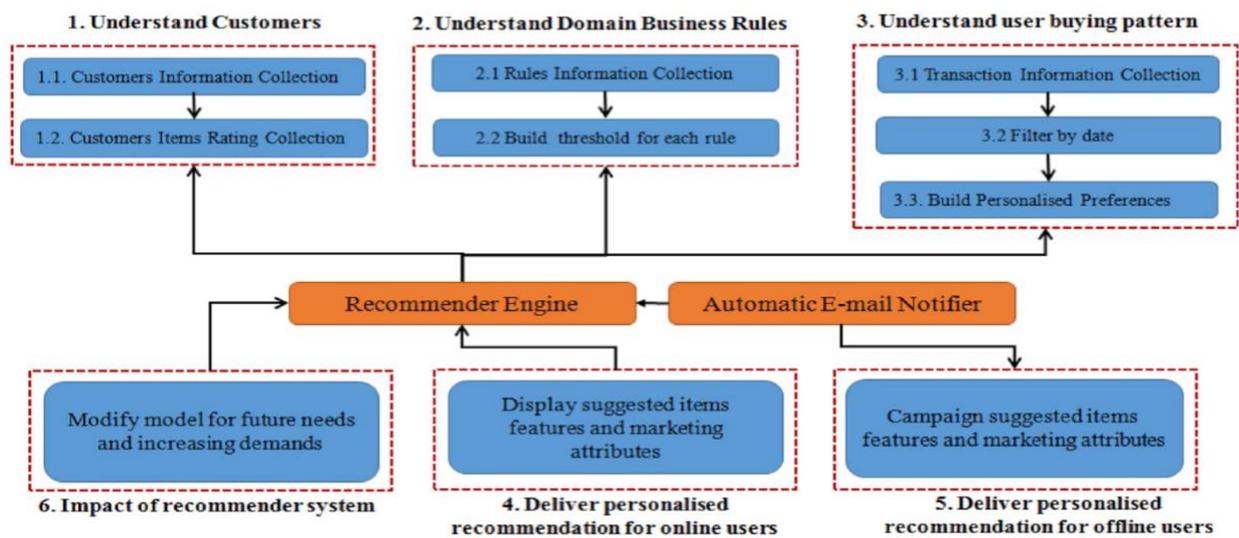


Personalized digital marketing recommender engine Rajat Kumar Behera, Angappa Gunasekaran, Shivam Gupta, Shampy Kamboj, Pradip Kumar Bala 2019

I cluster dei consumatori tengono conto solo degli individui che sono registrati al sito di e-commerce, e hanno già acquistato prodotti, non tenendo dunque conto dei *cold-start*, iscritti ma senza avere mai acquistato nulla, e dei *window-shopping costumers*, non iscritti e senza nessun acquisto effettuato. Infine analizzando il funzionamento del processo di raccomandazione di questa teoria, è possibile dividerlo in 6 step, interconnessi come mostrato nella Figura 11, dove i primi tre step si concentrano sull'analisi da parte dell'algorithmo dei consumatori, in base ai cluster definiti precedentemente, alla analisi delle regole di dominio del business ed infine a studiare il comportamento e il processo di acquisto del consumatore. Queste tre componenti vanno a definire il sistema di raccomandazione prima della sua applicazione, mentre invece lo step 4 consiste nella applicazione pratica del sistema di raccomandazione per gli utenti online, laddove invece nello step 5 il modello è implementato anche nelle strategie di Marketing rivolte agli utenti offline, attraverso un sistema di notifica tramite e-mail automaticamente gestito dall'algorithmo, come mostra la Figura

10. Nello step 6 vengono applicate le modifiche necessarie richieste dal contesto e dal cambiamento delle preferenze degli utenti in rete e offline, che permettono al sistema di rimanere competitivo rispetto allo sviluppo delle altre imprese e ai cambiamenti dell'ambiente esterno di riferimento. Infine, le metriche utilizzate per verificare i risultati ottenuti da questo sistema, analisi utile per apportare le giuste modifiche necessarie dopo il sesto step, riguardano l'efficacia e l'efficienza: la prima è misurata attraverso il monitoraggio delle revenues e del valore medio dell'ordine, mentre invece la seconda è misurata attraverso l'User Based Collaborative Filtering e l'Item Based Collaborative Filtering, cioè attraverso i dati riguardanti l'efficienza della classificazione dei diversi consumatori e della tassonomia di prodotto.

Figura 10: I 6 step del modello di raccomandazione



Personalized digital marketing recommender engine Rajat Kumar Behera, Angappa

Gunasekaran, Shivam Gupta, Shampy Kamboj, Pradip Kumar Bala 2019

2.3.1 Vantaggi e svantaggi dei *recommendation systems*

I sistemi di raccomandazione comportano numerosi vantaggi per aziende e consumatori, ma d'altra parte presentano anche numerosi svantaggi. Iniziando però ad analizzare tutti i benefici che l'adozione di algoritmi di raccomandazione comportano per i consumatori e le imprese, non si può non partire dalla percezione del consumatore, che attraverso tali sistemi ha una offerta il più possibile personalizzata ai suoi gusti e alle sue abitudini, aumentando di conseguenza anche la fiducia riposta da questo nell'impresa, la cui offerta è percepita come perfettamente aderente alle sue necessità. Inoltre i sistemi di raccomandazione permettono alle imprese di internalizzare una forma di

comunicazione che ha l'obiettivo di suggerire prodotti e servizi agli utenti, proprio come accade attraverso WOM ed e-WOM, ma rendendo il passaggio informativo un processo che l'utente percepisce come non forzato, come invece risulta essere il tentativo di controllo dei canali su cui avviene il passaparola, con la conseguenza di una riduzione della fiducia riposta nel brand. Da parte del consumatore inoltre risulta esserci un vantaggio riguardante il tempo necessario per trovare il prodotto ricercato, poiché, come ad esempio accade con i sistemi di raccomandazione usati da Netflix, analizzati nel terzo capitolo, di sicuro gli utenti avendo una offerta personalizzata impiegano molte meno risorse in termini di tempo per trovare il prodotto di cui hanno bisogno. Da parte delle imprese, i sistemi di raccomandazione aumentano anche gli acquisti non preventivati dai consumatori e dunque le *revenues*: questi sono i casi in cui un individuo che sta navigando sulla piattaforma di e-commerce di un brand alla ricerca di un prodotto, finisce per acquistarne anche un altro che non aveva considerato, grazie all'efficienza che tali sistemi hanno nel proporre items affini alle necessità anche inconsce dei consumatori. Inoltre, attraverso l'analisi dei dati utile a generare i sistemi di raccomandazione, le imprese riescono anche a comprendere con anticipo verso quale direzione si stanno rivolgendo le preferenze dei consumatori, in modo da utilizzare questa conoscenza nello sviluppo di nuovi prodotti o servizi quanto più aderenti possibile alle necessità mostrate dagli utenti.

Tra i principali fattori negativi dei *recommendation systems* troviamo la grande quantità di dati che questi sistemi necessitano per garantire una maggiore precisione nella personalizzazione dell'offerta per ogni utente, che si declina diversamente, come verrà analizzato nel prossimo paragrafo, a seconda che il sistema sia *content based* o sia *collaborative filtering*. Inoltre dal lato dei consumatori un elemento di negatività riguarda la privacy: molto spesso purtroppo sentiamo di casi di gestione non corretta dei dati sensibili da parte delle imprese, e questo può causare più cautela da parte degli individui nell'acconsentire al trattamento dei propri dati personali, provocando una riduzione della quantità di dati a disposizione delle imprese, che come visto è un fattore determinante nell'efficienza e nell'efficacia dei sistemi. Un ulteriore svantaggio è rappresentato dai sistemi, come il *collaborative filtering*, che richiedono una partecipazione attiva dell'utente che ad esempio deve dare un giudizio sui prodotti per collaborare nella creazione di una offerta a lui personalizzata. In merito è possibile fornire l'esempio di Netflix, che specialmente agli albori della sua storia, quando ancora i film erano distribuiti tramite posta, svolgeva una forma rudimentale di raccomandazione basata su un sistema di votazione dei contenuti con una scala da una a cinque stelle, per cui risultava difficile profilare un utente se queste non votava i film che aveva noleggiato. Infine un ultimo limite dei sistemi di raccomandazione è rappresentato dal *cold start problem*, ovvero la difficoltà che tutti i sistemi hanno

nell'implementare una personalizzazione dell'offerta quando si hanno nuovi prodotti o nuovi utenti sulla piattaforma

2.3.1.1 Pro e contro nei sistemi *content based* e *collaborative filtering*

Nell'analizzare più nello specifico i vantaggi e gli svantaggi dei sistemi di raccomandazione, è necessario richiamare la classificazione effettuata nel precedente paragrafo tra sistemi *content based* e sistemi basati sul *collaborative filtering*. Questi ultimi difatti, differiscono dai primi poichè sono sistemi flessibili e che dunque hanno come grande vantaggio la loro adattabilità in contesti differenti, in termini di quantità e tipologia di dati a disposizione. Un esempio di questo vantaggio sta nella possibilità di classificare le preferenze di un individuo non attraverso le caratteristiche degli item che ha acquistato, o che ha salvato tra i preferiti, bensì attraverso un sistema di voti che attribuisce un determinato livello di gradimento di un prodotto, permettendo di incrociare queste preferenze con quelle degli altri consumatori, in modo da formare cluster di preferenze. Questo permette di superare il problema dell'entrata di nuovi utenti, o di consumatori che non hanno ancora effettuato nessun acquisto sulla piattaforma, poiché sarà necessario immagazzinare le votazioni che questi danno ai prodotti, e molto spesso questi giudizi per new users sono ottenuti con brevi questionari in cui l'utente deve dare un voto a degli item causalmente selezionati dalla piattaforma, in modo da profilare il nuovo cliente. D'altra parte, lo svantaggio dei sistemi *content based* sta nella difficoltà di fornire una raccomandazione dettagliata e personalizzata dell'offerta ad un cliente definito, specie quando questo individuo è da poco presente sulla piattaforma, a causa della mancanza di dati sui contenuti acquistati. Viceversa nel caso dei *cold start items*, il sistema basato sul *collaborative filtering* fatica maggiormente nel classificare le preferenze su questo nuovo prodotto, a causa della mancanza di giudizi espressi dagli utenti. Questo comporta molto spesso che i nuovi prodotti risultano svantaggiati nelle fasi che seguono il lancio, poiché prima di entrare nelle classificazioni delle preferenze dei consumatori hanno bisogno di una grande quantità di voti. Per questa ragione un brand che tende a lanciare molti prodotti in un lasso breve di tempo, troverà in questa topologia di sistema uno svantaggio più accentuato rispetto alle imprese che invece hanno un intervallo di tempo tra il lancio di nuovi prodotti più dilatato. Nel caso dei sistemi *content based* invece, un prodotto da poco inserito sulla piattaforma non crea problemi nella raccomandazione, poiché basandosi sulle caratteristiche del prodotto stesso, associa queste a quelle dei prodotti già presenti, riuscendo in tal modo a predire le preferenze degli individui.

CAPITOLO 3

3.1 Netflix *recommendation system*

Netflix è la piattaforma che per prima ha unito la tecnologia, Internet e il mondo cinematografico, e ciò è stato reso possibile grazie ad un costante vantaggio sulla pur agguerrita concorrenza conquistato con un innovativo sistema di raccomandazione che si compone di diversi algoritmi. Questo risulta cruciale nell'offerta dei propri contenuti considerando che la scelta del consumatore è il punto intorno al quale deve costruirsi una raccomandazione efficace che permetta di semplificare questo processo, in modo da amplificare la soddisfazione dell'utente, il quale secondo Netflix⁴¹ dopo soli 60/90 secondi perde interesse avendo già visto le copertine e i titoli di circa 10/20 contenuti. Attraverso i *recommendation systems*, l'utente oltre a guardare un film, una serie TV o un documentario che gradirà particolarmente, lo avrà scelto in poco tempo grazie agli algoritmi che personalizzano il layout della pagina home di Netflix, semplificando anche visivamente le fasi di selezione del contenuto da guardare

3.1.1 L'alba dei sistemi di raccomandazione

Iniziando dai primi passi mossi da Netflix nei sistemi di raccomandazione, quando questa piattaforma aveva ancora la maggiore parte dei film distribuiti tramite i DVD per posta, e non con l'attuale piattaforma digitale, la società aveva concentrato i suoi sforzi per migliorare l'offerta ad ogni singolo utente cercando di implementare un sistema di rating basato su un punteggio dato dal consumatore ad ogni contenuto visto da una a cinque stelle. Nel corso degli anni poi con il passaggio al digitale e lo sviluppo della piattaforma online, il management ha sostenuto lo sviluppo di una raccomandazione sempre più efficiente anche attraverso l'*open innovation*, costituendo nel 2009 il *Netflix prize*⁴², ovvero una competizione a squadre aperta a chiunque con un premio per chi fosse riuscito ad aumentare la accuratezza del sistema di più del 10%, con lo scopo di ottenere un miglioramento di quello che era l'attuale algoritmo di raccomandazione, chiamato The Cinematch, il quale si basava

⁴¹ The Netflix Recommender System: Algorithms, Business Value, and Innovation Carlos A. Gomez-Uribe and Neil Hunt

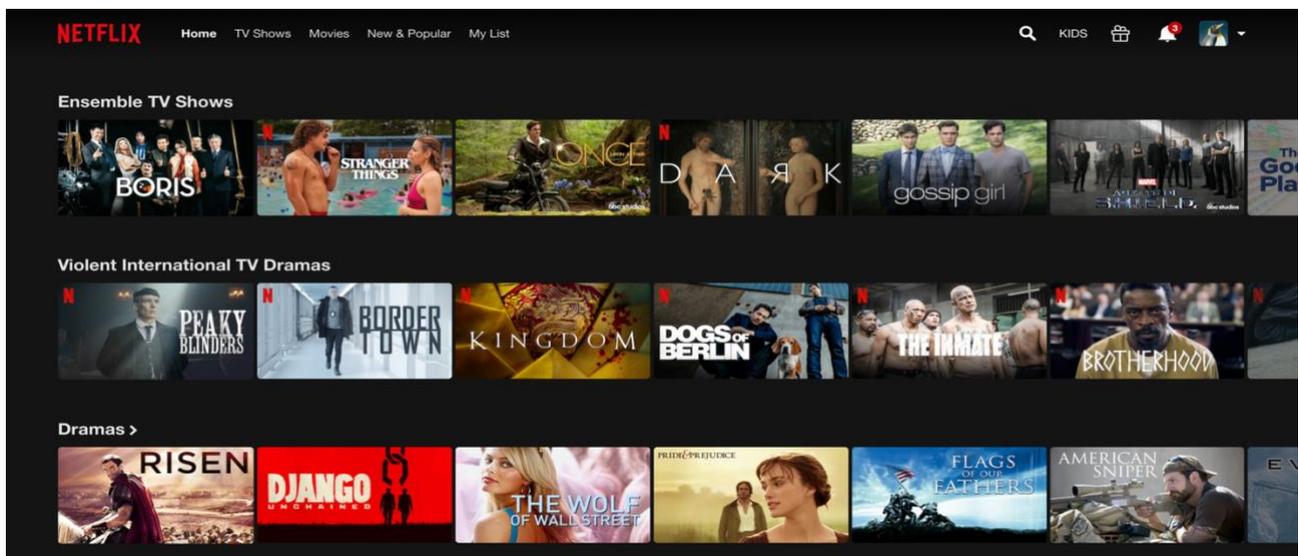
⁴² The Netflix prize, James Bennett e Stan Lanning

su un sistema di regressione lineare attraverso il quale personalizzava l'offerta in base ai rating dati dagli utenti ai contenuti visti.

3.1.2 Gli algoritmi alla base dei suggerimenti di Netflix

Come anticipato il *recomendation system* di Netflix si costituisce di numerosi algoritmi, che vanno a formare il ventaglio di contenuti mostrati nella homepage, il quale risulta essere il modo attraverso cui sono scelti l'80% dei contenuti, in uno schema composto da 40 righe per 75 film, serie o documentari l'una, e che implementano una sistema di ricerca personalizzata quando l'utente ricerca attivamente un contenuto tramite parole chiave, che risulta generare il restante 20% di contenuti visti sulla piattaforma. Gli algoritmi che saranno descritti sono il *Personalized Video Ranker*, il *Top-N Video Ranker*, *Tranding Now*, *Continue Watching* e infine il *Video-Video Similarity*: Il *Personalized Video Ranker*, come facilmente intuibile, basa il suo algoritmo sulla generazione di una serie di righe divise per genere personalizzate per ogni utente in base ai contenuti guardati fino a quel momento e ai rating dati. Come mostra la Figura 11, la riga composta da *Violent International TV Dramas* ha due livelli di raccomandazione: il primo livello si basa sul genere, mentre il secondo livello rende la proposta dei contenuti di questa riga unica per ogni utente.

Figura 11: Netflix homepage



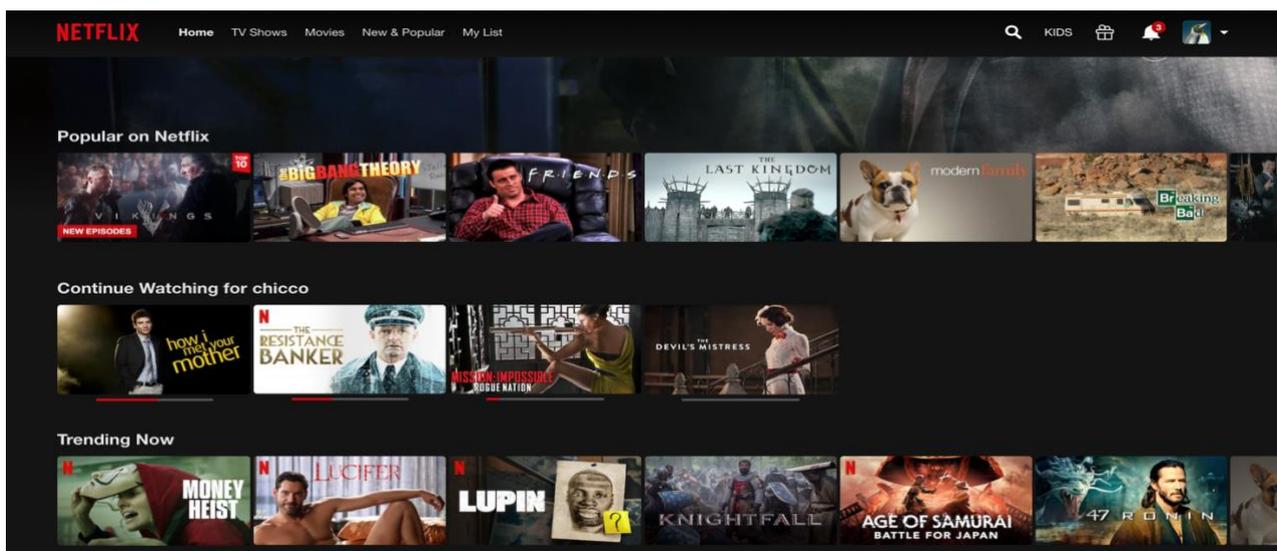
Netflix.com 2021

L'algoritmo *Top-N Video Ranker*, che ha un grado di libertà in più rispetto al PVR, perché non si basa sulla selezione di un catalogo ristretto su un particolare genere, ma valuta cataloghi più ampi scegliendo i top in assoluto considerando ad esempio l'area geografica dell'utente, e sfruttando sia dati personali, che la popolarità dei contenuti.

Il *Trending Now* invece va a comporre i film suggeriti nella omonima riga mostrata nella Figura 13, basandosi principalmente sulle due categorie: da una parte vengono proposti tutti i generi di film definiti stagionali, incrociati con le preferenze dell'utente che come si può notare costituisce la base imprescindibile per ogni declinazione degli algoritmi di Netflix, ovvero quei contenuti che sono legati a particolari periodi dell'anno come i film natalizi per il Natale o i film Horror per il periodo di Halloween; d'altra parte questo algoritmo propone quei film considerati un *Trending Now one-off*, a causa di un particolare evento che non si ripete con una scadenza regolare come ad esempio un terremoto o una grande operazione internazionale contro il narcotraffico.

Il *Continue Watching* invece si basa sull'analizzare il comportamento che il consumatore ha assunto durante la fruizione di un contenuto che non ha terminato di vedere, e proponendo in questa riga mostrata nella Figura 12 quei contenuti che secondo l'algoritmo hanno ancora un potenziale di attrazione per l'utente, considerando ad esempio il punto di abbandono del film o della serie TV oppure gli altri contenuti guardati dopo questo presente nel *Continue Watching*.

Figura 12: Netflix homepage



Netflix.com 2021

Infine, il *Video-Video Similarity* crea la riga del *Because You Watched*, agendo su due livelli, il primo della scelta dei video in base a quelli già visti, ed il secondo scegliendo quali generi in base ai gusti dell'utente, il quale risulta essere più personalizzato rispetto al primo.

Come anticipato all'inizio del paragrafo sono implementati anche algoritmi che non vanno specificatamente a comporre i suggerimenti proposti nella homepage, ma che più in generale amplificano il piacere dell'utilizzo della piattaforma, come ad esempio l'algoritmo dell'*Evidence Selection*, che si occupa di scegliere quale tipo di informazione comunicare all'utente in modo sintetico su un film, se gli attori, eventuali premi vinti o il regista, oppure quale immagine usare come copertina del contenuto, oppure come un algoritmo in grado di capire se saltare o meno il riassunto

delle puntate precedenti di una serie presente in un episodio, comparando il tempo intercorso tra l'ultimo episodio visto ed il nuovo, con la velocità media dell'utente nel guardare una serie TV, fenomeno che nel caso estremo di visione di più episodi consecutivamente prende il nome di binge-watching. Infine, l'ultima categoria di algoritmi sono quelli che costituiscono la personalizzazione del sistema di ricerca, poiché molto spesso gli utenti non usano il titolo del film bensì una parola chiave o un concetto ad esso correlato, e lo scopo degli algoritmi del sistema di ricerca è proprio quello di suggerire i contenuti che sono più aderenti alla parola ricercata. Inoltre, considerando la concreta ipotesi che un utente possa cercare un film non presente su Netflix, l'algoritmo di ricerca deve rendere questa mancanza il meno pesante possibile nella percezione del consumatore, offrendo una ampia gamma di alternative al film ricercato ma non presente nell'offerta, lavorando per analogia di genere, cast e regista, dati come sempre incrociati con le preferenze registrate dell'utente.

3.2 Zalando recommendation system

Zalando è una azienda nata nel 2008, definita anche “*e-tailer*” come crasi di e-commerce retailer, che oggi conta 15.000 dipendenti, attiva in 17 paesi con 27 milioni di consumatori che hanno generato nel 2020 ricavi per 4,5 miliardi di euro⁴³. Per sopportare la ingente mole di dati che una società come questa necessariamente deve utilizzare nei propri processi per aumentare il valore dell'offerta, Zalando ha dal 2016 avuto accesso alla piattaforma BigQuery di Google Analytics 360 Suits, uno strumento che permette l'analisi dei dati raccolti dall'impresa. Una volta analizzate, queste informazioni costituiscono il carburante dei sistemi di raccomandazione utilizzati dall'e-commerce tedesco, che risultano avere due obiettivi principali, considerando da un lato la necessità di aumentare la percezione di valore da parte del consumatore, migliorando KPIs come l'engagement ed il *retention rate*, e d'altra parte la necessità interna dell'impresa di agire ottimizzando la profittabilità della tecnologia, che si traduce in una efficienza operativa, misurata con tre variabili, ovvero la adattabilità della tecnologia implementata, il costo e la complessità della manutenzione nel medio-lungo periodo, ed infine la rapidità di integrazione nel sistema dei nuovi dati continuamente disponibili.

⁴³ <https://cloud.google.com/customers/zalando?hl=it>

3.2.1 L'architettura del sistema

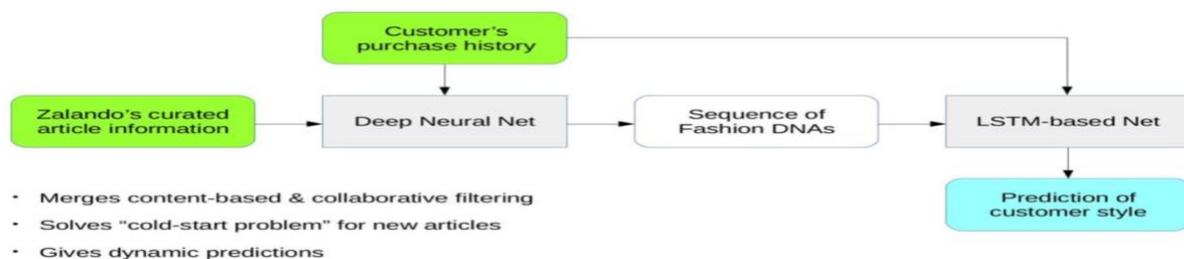
I sistemi di raccomandazione analizzati nel prossimo paragrafo compongono il ventaglio degli strumenti utilizzati da Zalando affinché i dati siano la guida strategica dell'impresa con lo scopo di far nascere da questi una forma di miglioramento dell'esperienza del consumatore, suggerendo i prodotti in base alle sue preferenze cercando di replicare il passaggio informativo dell'e-WOM, ma internalizzandolo. L'architettura dei sistemi di raccomandazione può essere divisa in due macro-gruppi, considerando da un lato il lavoro svolto offline, e dall'altro le operazioni svolte online. Le operazioni svolte offline, da considerarsi propedeutiche alla personalizzazione live che avviene tramite le operazioni online, iniziano registrando il comportamento del consumatore sulla piattaforma, raggruppando per ogni utente una serie di azioni chiave che questo svolge, permettendo in tal modo di semplificare i successivi passaggi, e più precisamente rendendo la futura ricerca di dati su un cliente più semplice. Una volta raccolti e categorizzati, questi dati risulteranno utili per due processi successivi, ovvero l'analisi delle caratteristiche degli articoli da un lato, e l'analisi del consumatore dall'altro, considerando che per l'articolo sono valutati il numero di click o il numero di acquisti in un determinato lasso di tempo, variabili utili anche per il secondo processo, che terrà conto del tempo passato sul sito o della predilezione per un brand in particolare del cliente. Il processo offline si conclude con una funzione ottenuta grazie al sistema Learn to rank, il quale darà un punteggio sia all'articolo che all'utente, che sarà determinante nell'architettura online dei sistemi di raccomandazione. Il secondo gruppo di operazioni dell'architettura dei sistemi di raccomandazione che fanno parte del centro gravitazionale della personalizzazione di massa di Zalando sono le operazioni online. Lo scopo principale di questo insieme di processi è quello di ottenere una personalizzazione live, ottenuta grazie ad un software come Apache Solr, che permette in soli 5 millisecondi di interrogare il set informativo generato dalle operazioni offline. Da questo poi è possibile generare un ranking dei prodotti da raccomandare tramite un motore di ricerca che interagisce con un Reco servlet, ovvero un oggetto scritto in linguaggio Java, che è il vero artefice della selezione dal ranking dei prodotti raccomandabili dell'offerta personalizzata per ogni utente.

3.2.2 I modelli di raccomandazione utilizzati da Zalando

Ma in che modo Zalando, in quanto *data-driven company*, riesce ad ottenere informazioni utili da questi dati? Il primo sistema di raccomandazione si basa sui prodotti simili a quelli ricercati dal consumatore, e dunque non utilizza necessariamente le preferenze storiche del consumatore, bensì parte dal bisogno attuale mostrato con la ricerca. In questo contesto anche i prodotti suggeriti nella sezione "*similar items*" possono però essere raccomandati incrociando i dati delle precedenti ricerche e soprattutto degli acquisti già effettuati, implementando un sistema che suggerisce non solo articoli

simili, bensì anche suggerimenti su come completare l'outfit partendo dal prodotto ricercato. Un secondo sistema di raccomandazione usato nella piattaforma, che basa maggiormente i suoi risultati sulla analisi della serie storica di dati acquisiti sul cliente, è la sezione “*recommended for you*” che opera esclusivamente considerando i brand più acquistati dall'utente e poi partendo da questi seleziona articoli in base alle ricerche recentemente effettuate. Un altro sistema è l'Algorithmic Fashion Companion, il quale incrociando i dati del consumatore, come i prodotti inseriti nel carrello o quelli precedentemente acquistati, con i giudizi di *fashion stylists* che lavorano con l'impresa, compone un sistema di raccomandazione di outfit che rappresenta il 40% delle scelte degli acquisti che i consumatori compiono dai 2000 marchi offerti sulla piattaforma⁴⁴. I sistemi di raccomandazione dell'*e-tailer* utilizzano inoltre metodi di *click chain analysis*, ovvero di immagazzinamento della sequenza dei click di ogni consumatore in modo da studiare il suo comportamento e personalizzare l'offerta secondo le esigenze mostrate dall'utente. Questo sistema è inserito all'interno di un modello di raccomandazione di prodotti definito Long-Short Term Memory⁴⁵ e schematizzato nella Figura 13, il quale ha lo scopo di predire non solo nel breve termine le preferenze del consumatore, ma di comprendere inoltre il modo in cui queste variano nel lungo periodo. Inoltre, questo sistema di raccomandazione permette di superare il problema dei *cold-start items*, ovvero di quei prodotti che sono stati inseriti da poco nella gamma della piattaforma e di cui non si ha una storia di acquisto da parte dei consumatori, tipico di sistemi come il *collaborative filtering*, basati sui giudizi da parte del consumatore riguardo ai prodotti. Questo è possibile grazie all'utilizzo del Fashion DNA⁴⁶, ovvero un sistema che attraverso un vettore per ogni immagine permette di rappresentare una distribuzione geometrico-spaziale dei prodotti, considerando le direzioni con diverse variabili come la fantasia, il colore, il tipo di prodotto o il marchio, come rappresentato dall'immagine sottostante. Questo sistema permette così di predire il comportamento e la conseguente scelta del consumatore senza necessariamente sapere il giudizio dell'utente sull'articolo specifico, bensì agendo per analogia rispetto alle preferenze già mostrate.

Figura 13: Long-Short Term Memory



⁴⁴ <https://www.forbes.com/sites/bernardmarr/2019/09/20/the-amazing-ways-retail-giant-zalando-is-using-artificial-intelligence/?sh=6bb5dbc94d93>

⁴⁵ <https://research.zalando.com/welcome/mission/research-projects/forecasting-the-customers-preference/>

⁴⁶ https://research.zalando.com/welcome/mission/research-projects/improving-fashion-item-encoding-and-retrieval/?gh_src=4n3gxh1

I problemi principali che i sistemi di raccomandazione guidati dall'AI e dal Machine Learning di Zalando sono costituiti dal fatto che da un lato l'analisi dei dati, e l'anticipo della scelta del consumatore, impiega ancora troppo tempo ad adattarsi alla velocità con cui una moda prende piede sul web grazie ai post degli *influencer*, e dall'altro dalla possibilità che non sempre il profilo di un consumatore è utilizzato unicamente da questo, ma può ad esempio essere un unico profilo condiviso da una coppia, così da rendere necessaria l'implementazione di un sistema che capisca, nei primi minuti di utilizzo della piattaforma, l'identità del consumatore in modo da migliorare la personalizzazione dell'offerta. Se la seconda problematica non ha ancora avuto una soluzione, per il primo problema i sistemi di raccomandazione sono sostenuti da altri strumenti come i già citati *stylist* che lavorano ogni istante per trovare gli outfit del momento, anche seguendo l'attività social degli *influencer*, o da strumenti come il *deep learning visual research*, il quale permette di cercare degli indumenti che l'utente ha visto ad esempio su un post, inserendo l'immagine in questa funzione di ricerca.

CONCLUSIONE

Il percorso concettuale svolto in questo lavoro ha come risultato finale la grande attenzione che le imprese devono porre nell'implementazione dei dati e nel loro utilizzo all'interno dei processi delle marketing unit. Il WOM, e con la *digital transformation* in modo ancora più marcato l'e-WOM, sono modelli di comunicazione tra consumatori, in cui avvengono scambi di giudizi che influenzano le decisioni di acquisto. Come è stato possibile notare però, le imprese non possono controllare le informazioni che transitano, né possono modellare a loro piacimento i canali su cui queste vengono scambiate, poiché tale azione avrebbe un rischio molto alto di far cadere la fiducia che il consumatore ha nei confronti del brand. Per questo motivo la tesi argomentata in questo lavoro è quella di proporre il sistema di raccomandazione come una forma di passaparola indiretto, cioè difficilmente percepito dal consumatore come tale, ed internalizzato nelle strategie dell'impresa, attraverso il quale il brand riesce a guidare le preferenze dei propri utenti, proponendo una personalizzazione di massa dell'offerta. Lo sviluppo di questi sistemi è reso possibile dagli esponenziali passi in avanti che la tecnologia ha avuto nella gestione, analisi ed elaborazione dei dati, con strumenti legati all'intelligenza artificiale come il Machine Learning, che permettono di avere una gestione ottimizzata delle risorse a disposizione. Queste risorse sono rappresentate proprio dai Big Data, così definiti per le loro dimensioni e soprattutto per la immensa velocità con la quale aumentano ogni anno, come mostrato dall'esempio di Rasetti riportato nell'introduzione. I sistemi di raccomandazione, dunque, rappresentano la nuova frontiera dell'offerta delle imprese, uno sviluppo reso possibile anche dal fatto che oramai la comunicazione non è più una cosa esclusivamente tra uomini, ma è e sarà sempre più un qualcosa tra uomini e macchine, ma soprattutto tra sole macchine, automatizzando la gestione e la elaborazione dei dati per aumentare l'efficienza delle imprese, e di conseguenza la fidelizzazione dei consumatori.

Le imprese che non stanno rivolgendo la bussola delle loro strategie verso questa direzione, a mio modesto parere, non avranno un futuro roseo, ma si scontreranno con la rapidità delle imprese *tech-oriented*, di comprendere le sfumature più labili dei bisogni di ogni singolo individuo per trasformarle in prodotti presenti sul mercato, andando sempre con maggiore efficacia a predire i bisogni stessi prima che i consumatori ne siano consci.

Il futuro forse non è più così hollywoodiano e le imprese devono attivarsi per plasmare la realtà in un film di fantascienza.

BIBLIOGRAFIA:

- Antonino Freno, Practical Lessons from Developing a Large-Scale Recommender System at Zalando
- Anubhav Mishra and Satish S. M., E-WOM: Extant Research Review and Future Research Avenues
- Carlos A. Gomez-Uribe and Neil Hunt, Netflix, Inc. The Netflix recommender system: algorithms, business value, and innovation
- Carlos A. Gomez-Uribe and Neil Hunt, The Netflix Recommender System: Algorithms, Business Value, and Innovation
- Christy M.K. Cheung, Dimple R. Thadani The impact of electronic word-of-mouth communication: A literature analysis and integrative model
- Jiska Eelen, Peren Özturan, Peeter W.J. Verlegh, The differential impact of brand loyalty on traditional and online word of mouth: The moderating roles of self-brand connection and the desire to help the brand
- Liye Ma, Baohong Sun, Machine learning and AI in marketing – Connecting computing power to human insights
- Lun-ping Hung, A personalized recommendation system based on product taxonomy for one-to-one marketing online
- M.Rasetti, “Il dado è tratto. Big data e IA tra scienza e società”
- Manuela López and María Sicilia, Determinants of E-WOM Influence: The Role of Consumers’ Internet Experience
- Matteo De Angelis, Vito Tassiello, Cesare Amatulli, Michele Costabile, How language abstractness affects service referral persuasiveness
- Nobuyuki Fukawa, Sunil Erevelles, Linda Swayne, Big Data consumer analytics and the transformation of marketing
- Noraini Sa’ait, Agnes Kanyan and Mohamad Fitri Nazrin, The Effect of E-WOM on Customer Purchase Intention
- Nuria Huete-Alcocer, A Literature Review of Word of Mouth and Electronic Word of Mouth: Implications for Consumer Behavior
- On Braggarts and Gossips: Why Consumers Generate Positive but Transmit Negative Word-of-Mouth

- Rajat Kumar Behera, Angappa Gunasekaran, Shivam Gupta, Shampy Kamboj, Pradip Kumar Bala, Personalized digital marketing recommender engine
- Robert M. Bell e Yehuda Koren, Lessons from the Netflix prize challenge
- Sanjeev Verma & Neha Yadav, Past, Present, and Future of Electronic Word of Mouth (EWOM)
- Shaphali Gupta & Agata Leszkiewicz & V. Kumar & Tammo Bijmolt & Dmitriy Potapov, Digital Analytics: Modeling for Insights and New Methods
- Shoshanna Zuboff, Il capitalismo della sorveglianza

SITOGRAFIA:

- https://blog.osservatori.net/it_it/big-data-e-tutela-dei-dati-personali-%C3%A8-possibile
- <https://cloud.google.com/customers/zalando>
- <https://consumerresearcher.com/jonah-berger-contagious>
- <https://enricomtomassi.com/passaparola-marketing/>
- <https://enricomtomassi.com/zmot-social-conversione/>
- <https://it.businessinsider.com/ascoltate-sempre-i-vostri-clienti-la-fiducia-dei-consumatori-si-basa-su-recensioni-aperte-e-trasparenti/>
- <https://it.businessinsider.com/che-cosa-sono-davvero-i-big-data-chi-ci-guadagna-e-come-fa-e-come-siamo-messi-in-italia/>
- <https://research.zalando.com/welcome/mission/research-projects/forecasting-the-customers-preference/>
- https://research.zalando.com/welcome/mission/research-projects/improving-fashion-item-encoding-and-retrieval/?gh_src=4n3gxh1
- <https://tech.everyeye.it/notizie/nomofobia-fenomeno-in-crescita-smartphone-diventa-estensione-del-corpo-301934.html>
- <https://towardsdatascience.com/zalando-dress-recomendation-and-tagging-f38e1cbfc4a9>
- <https://wearesocial.com/digital-2021>
- <https://www.bigdata4innovation.it/big-data/come-i-big-data-vengono-utilizzati-nel-marketing-digitale/>
- <https://www.digital4.biz/marketing/big-data-e-analytics/big-data-cosa-sono-e-perche-grazie-alle-analitiche-il-business-continua-a-crescere/>
- <https://www.forbes.com/sites/bernardmarr/2019/09/20/the-amazing-ways-retail-giant-zalando-is-using-artificial-intelligence/>
- <https://www.freeway.it/analytics/marketing-operativo-data-driven>
- <https://www.ilsole24ore.com/art/big-data-noi-siamo-quello-che-compriamo-e-supermercato-sa-meglio-noi-ADCs0RJB>
- <https://www.ilsole24ore.com/art/in-italia-solo-un-impresa-4-sfrutta-bene-tesoro-dati-ADi3uY3>
- <https://www.infodata.ilsole24ore.com/2017/05/04/quanto-vale-mercato-europeo-dei-dati/>
- <https://www.insidemarketing.it/glossario/definizione/viral-marketing/#:~:text=Significato%20di%20Viral%20marketing,cerchie%20opinioni%2C%20consigli%2C%20notizie%20originali>

- <https://www.insidemarketing.it/glossario/definizione/word-of-mouth/>
- <https://www.zerounoweb.it/analytics/big-data-come-produrre-insight-per-una-strategia-marketing-data-driven/>