



UNIVERSITÀ POLITECNICA DELLE MARCHE

FACOLTÀ DI ECONOMIA “GIORGIO FUÀ”

Corso di Laurea Triennale in Economia e Commercio

Big data marketing

Big data marketing

Relatore

Prof. Federica Pascucci

Rapporto Finale di:

Paesani Davide

Anno Accademico 2020/2021

INDICE

INTRODUZIONE	pag. 4
CAPITOLO 1: I BIG DATA	pag. 6
1.1 Big Data: caratteristiche, evoluzione e fonti	pag. 6
1.2 Una nuova figura professionale: il data scientist	pag. 10
1.3 La data economy	pag. 14
1.4 La profilazione del consumatore: tra libertà, canalizzazione della scelta e tutela della privacy	pag. 16
CAPITOLO 2: BIG DATA MARKETING	pag. 19
2.1 Data-driven marketing: l'esigenza di un approccio basato sui dati	pag. 19
2.2 Big Data analytics: un nuovo strumento per il supporto alle decisioni	pag. 23
2.3 Big Data analytics in una prospettiva di marketing mix	pag. 30
2.4 Big Data analytics nel marketing: applicazioni	pag. 36
CAPITOLO 3: UTILIZZO DEI BIG DATA NEL MARKETING: ALCUNI CASI ESEMPLIFICATIVI	pag. 45
3.1 Facebook	pag. 45
3.2 Google	pag. 47
3.3 Netflix	pag. 49
3.4 Amazon	pag. 51
3.5 Spotify	pag. 54
CONCLUSIONI	pag. 57
BIBLIOGRAFIA	pag. 59
SITOGRAFIA	pag. 60

INDICE DELLE FIGURE

Figura 1: Il processo di creazione di un vantaggio competitivo sostenibile	pag. 28
Figura 2: Il panorama delle Big Data analytics visto attraverso il modello del marketing mix	pag. 31
Figura 3: L'ampiezza diagnostica delle Big Data analytics nel marketing	pag. 36

INTRODUZIONE

Al giorno d'oggi ogni persona è immersa, senza accorgersene, in un oceano di dati. Basti pensare che ogni minuto vengono scambiate, nel mondo, milioni di informazioni: ogni 60 secondi, su Facebook, vengono creati 3.3 milioni di post, pubblicati 510.000 commenti e aggiornati 293.000 stati; su Twitter vengono inviati 470.000 tweet; su WhatsApp vengono scambiati 38 milioni di messaggi; su Google vengono effettuate 3.8 milioni di ricerche.

Inoltre, oggi l'ambiente circostante è in grado di "parlare" e di "ascoltare": le strade, le automobili, le case, i siti Web, i social, i negozi, i prodotti, sono tutti diventati connessi e sincronizzati. L'insieme di tutte queste nuove tecnologie permette di raccogliere sempre più dati dalle persone, sempre più diversi e sempre più velocemente.

Politica, economia, società: i Big Data penetrano in ogni contesto, sia pubblico che privato.

Per le imprese diventa indispensabile basare le proprie azioni su dati certi ed esaustivi: a differenza di molte mode tecnologiche, infatti, i Big Data non sono un trend ma una necessità gestionale. Questi nuovi datasets, qualora sfruttati correttamente, possono contribuire all'aumento della competitività, della crescita del business e dell'innovazione dell'impresa in diversi modi: aiutando a capire le reazioni dei mercati e le loro percezioni dei brand; identificando i fattori chiave che spingono le persone ad acquistare un certo prodotto o servizio e poi segmentandole, sulla base di tali fattori, per personalizzare quanto più possibile le strategie d'azione; abilitando nuovi tipi di sperimentazioni, consentite dalla disponibilità di dati inediti; guadagnando in

predittività, grazie ad un puntuale storico di informazioni ad ampio raggio, tale da consentire simulazioni molto accurate.

Nel presente elaborato viene sviluppata la tematica dei Big Data e l'impatto della "rivoluzione" che i Big Data hanno portato nel marketing.

Nel primo capitolo viene descritto che cosa sono nella pratica questi "nuovi dati", le loro fonti, l'impatto che hanno avuto in vari settori dell'economia e il loro rapporto con la privacy del consumatore.

Nel secondo capitolo viene affrontato il tema del "Big Data marketing", o marketing "data-driven", cioè guidato dai dati, come conseguenza dell'esigenza di sfruttare questa enorme mole di dati generati ogni secondo. In particolare, l'attenzione è posta sul concetto di "Big Data analytics", cioè i vari tipi di analisi che le aziende svolgono a partire dai dati, e del loro impatto nei processi aziendali.

Nel terzo capitolo vengono forniti esempi concreti di come alcune grandi aziende utilizzino i Big Data che i loro utenti generano più o meno inconsapevolmente per capirne e monitorarne il comportamento, al fine di fornire loro la migliore esperienza di acquisto possibile mediante un'accurata personalizzazione dei contenuti.

CAPITOLO 1: I BIG DATA

1.1 Big Data: caratteristiche, evoluzione e fonti

Oggi più che mai qualsiasi operazione che svolgiamo è influenzata da una fonte molto rilevante, l'informazione. Sin dall'antichità, seppur non disponendo degli strumenti tecnologici odierni che garantiscono un enorme flusso e varietà di informazioni, l'uomo ne ha sempre dato un gran peso, cercando sempre di avere una base solida e controprovata che motivasse le sue azioni.

Non abbiamo una definizione univoca di Big Data (Del Mastro e Nicita, 2019), ma possiamo basarci su tre caratteristiche basilari chiamate "tre V": Volume, Velocità e Varietà. La prima V fa riferimento all'immensa quantità di dati che vengono generati ogni secondo (circa $2,5 \times 10^{21}$ al giorno). La varietà si riferisce alla differente tipologia dei dati che vengono generati, accumulati ed utilizzati. Prima dell'epoca dei Big Data si prendevano in considerazione per le analisi principalmente dati strutturati¹ (hanno un formato e una lunghezza definiti) mediante l'uso di database relazionali, mentre oggi si prendono in considerazione anche dati non strutturati² (non seguono uno specifico formato), ad esempio immagini, video, audio, testi (ecc) e dati semistrutturati, ad esempio un atto notarile con frasi fisse e frasi variabili. Per ultimo la velocità, naturalmente intesa come la rapidità con cui i nuovi dati vengono generati. Sono importanti non solo la celerità nella generazione dei dati, ma anche la necessità che questi dati/informazioni arrivino in sistema real-time al fine di effettuare analisi su di essi.

¹ Provengono dal database CRM, dai sensori, dal GPS, ecc.

² Provengono dal Web, dai social media, dalle e-mail, ecc.

Queste tre caratteristiche hanno generato le odierne ed ulteriori due V dei Big Data: la veridicità e il valore. Se i dati alla base delle analisi sono poco accurati, i risultati delle analisi non saranno migliori. Visto che su tali risultati possono essere basate delle decisioni, è fondamentale assegnare un indice di veridicità ai dati su cui si basano le analisi, in modo da avere una misura dell'affidabilità. Il valore è riferito alla capacità di trasformare i dati in risultati. Un progetto Big Data necessita di investimenti, anche importanti, per la raccolta granulare dei dati e la loro analisi. Prima di avviare un'iniziativa è importante valutare e documentare quale sia il valore effettivo portato al business.

Oltre che per l'enorme quantità delle informazioni, è doveroso sottolineare che il termine "big" è utilizzato per definire le tecnologie usate per estrarre conoscenza e valore dai dati raccolti, che si discostano molto dai semplici software tradizionali disponibili sul mercato da molti decenni.

Per ultimo ma sicuramente non per importanza, va considerato il dato come bene economico (Delmastro e Nicita, 2019). È vero che stiamo parlando di una caratteristica connessa con i diritti della persona, ma l'analisi economica è una questione pragmatica che prescinde da qualsiasi valore etico e giuridico, poiché sfrutta economicamente le informazioni fornite per costruire o approcciarsi un business sopra. Il problema principale derivante da questa analisi riguarda l'ambiguità e incompletezza circa la proprietà degli stessi. La proprietà attribuisce un insieme di usi al soggetto titolare, ma nel rilasciare consensi possibili agli usi c'è una sottile linea tra l'alienabilità e la delega funzionale all'identificazione dell'utente per permettere l'uso del servizio.

Le prime testimonianze³ risalenti all'archiviazione ed utilizzo dei dati risalgono addirittura a fine 1800, quando un dipendente dell'ufficio di censimento USA inventò un sistema di classificazione dei dati relativi al censimento della popolazione, permettendo di svolgerlo interamente in 3 anni al posto dei consueti 10. Il fatto che più si avvicina a quanto definiamo oggi Big Data è datato 1965: in un data center americano viene creato un database relazionale che associa la colonna ID prodotto sia nella sezione "prodotti venduti" che in quella "informazioni prodotto".

Le analisi erano sì statiche e limitate, estraendo dati solo a consuntivo, ma lavorando sulle grandezze aggregate hanno dato la base per lo sviluppo delle stesse, e la creazione dei database relazionali verso l'inizio degli anni Ottanta. Proprio in questo periodo, l'analisi diventa dinamica perché l'estrazione avviene sia aggregata che in dettaglio, sui dati operazionali e previsionali basati su statistiche anziché solo a consuntivo.

La vera rivoluzione è grazie a Internet a fine anni '90, quando la rete permette di condividere con altri utenti a distanza le informazioni che vengono rilevate in un qualsiasi luogo, interno o esterno al creatore di dati. Ecco che in pochi decenni l'enorme disponibilità di fonti di dati ha fatto crescere in maniera esponenziale tutte le scienze e le tecniche ricadenti sotto l'ambito della digitalizzazione, tra cui possiamo citare:

- l'automazione, il passaggio di attività umane allo svolgimento automatico tramite tecnologie;
- l'intelligenza artificiale, rendere intelligenti le macchine e farle funzionare con prospetti medio-lungo termine nel proprio ambiente;

³ <https://www.themarketingfreaks.com/2019/11/big-data-cosa-sono-la-storia-le-caratteristiche-le-analisi-ed-esempi/>

- il data mining, l'estrazione di informazioni commercialmente rilevanti da una grande massa di dati;
- il machine learning (apprendimento automatico), la capacità di un computer di apprendere e modificare i propri processi in base alle informazioni acquisite;
- la robotica, la creazione di sistemi artificiali per sostituire i compiti o servizi per le persone.

Al giorno d'oggi, i Big Data presentano una grande varietà di fonti, che si dividono in due grandi macrocategorie: "fonti-uomo" e "fonti-macchina". I dati, infatti, si dividono generalmente in "human generated" (o "user-generated content", cioè contenuti generati dagli utenti) e "machine generated".

I primi provengono dalle piattaforme di social network (Facebook, LinkedIn), di blogging (Blogger, WordPress) e microblogging (Twitter, Tumblr), di social news (Digg, Reddit), di condivisione multimediale (Instagram, YouTube), da siti di domande e risposte (Yahoo Answers), da siti di recensioni (TripAdvisor), da portali di eCommerce (Amazon, eBay) e dai click stream da siti Web. I secondi provengono da sorgenti quali smartphones, sensori GPS (geolocalizzazione), dispositivi IoT (Internet of Things), RFID (Radio Frequency Identification), contatori delle utenze energetiche, sistemi di High Frequency Trading dei mercati finanziari, dispositivi biomedicali, macchinari industriali, ecc.

Un particolare "sottoinsieme" dei due tipi di dato sopracitati è quello dei c.d. dati "business generated", con cui si intendono tutti i dati, human o machine generated, prodotti internamente ad un'azienda, relativi alle attività data-driven dei processi di business della stessa. Molti di essi sono dati storici, memorizzati staticamente e

normalmente archiviati separatamente nei database relazionali della relativa area funzionale.

Infine, vanno anche menzionati gli “open data”: sono dati che possono essere utilizzati, riutilizzati e ridistribuiti da tutti gratuitamente, con l’unico vincolo di indicarne la fonte⁴ e di condividerli utilizzando la stessa licenza da questi adottata. Non essendoci uno standard definito, sono disponibili in diverse forme:

- Dati statici: sono disponibili in quei formati che non consentono un’extrapolazione immediata degli stessi (ad esempio GIF, JPG, BMP, PDF, Word);
- Dati strutturati: sono disponibili in formati che ne consentono l’elaborazione e la sistematizzazione in forma strutturata (ad esempio Excel o OpenOffice Calc);
- Dati presenti in database: sono presenti all’interno dei database gestiti dall’ente. Di norma, di questi dati è possibile scegliere il formato di esportazione più idoneo (ad esempio XML o altri formati aperti).

1.2 Una nuova figura professionale: il data scientist

Le opportunità offerte dalla rivoluzione dei Big Data, per essere colte appieno, richiedono l’introduzione di nuove competenze all’interno delle imprese: al giorno d’oggi sono richieste nuove figure professionali, in grado di coniugare capacità informatiche nell’utilizzo delle tecnologie innovative per la gestione dei dati, competenze di machine learning e intelligenza artificiale, conoscenza del business in cui opera l’azienda e abilità nel comunicare le analisi svolte al management aziendale.

⁴Ad esempio, il data.gov del governo americano, il CIA World Factbook o il portale Open Data Portal dell’Unione Europea.

Questo profilo multidisciplinare è quello del data scientist, una delle figure professionali oggi più richieste sul mercato.

Fino a qualche anno fa l'analisi dei dati era appannaggio degli statisti, che dovevano occuparsi di decidere su cosa raccogliere i dati, scegliere in che modo farlo, come analizzarli e come codificare i risultati. Con l'avvento di Internet, tuttavia, i dati a disposizione sono in sovrabbondanza: da qui la nascita della *data science*, definita come il "processo di rilevamento di insight non immediatamente evidenti da enormi volumi di dati strutturati e non, utilizzando metodi che comprendono la statistica, l'apprendimento automatico, il data mining e la predictive analytics" (Iandiorio, 2019). Il compito del data scientist va oltre quello del semplice statista: egli, infatti, deve anche interpretare i dati in chiave di marketing, quindi in termini predittivi, basandosi su comportamenti reali e non su ipotesi, con l'obiettivo di offrire una panoramica precisa dei problemi che l'impresa dovrà affrontare e risolvere. In poche parole, il data scientist deve trasformare i dati in informazioni e queste ultime in conoscenza.

Nel 2014 solo il 13% delle aziende possedeva al proprio interno profili con queste caratteristiche, mentre solo il 2% li riconosceva formalmente come data scientist. Nel 2015 si è registrata una notevole crescita: dal 13% si è passati al 30%, mentre dal 2% al 4%. Nel 2016 il tasso di aziende con un data scientist al loro interno è salito solo al 31%, tuttavia si è assistito a un significativo passo in avanti in termini di formalizzazione della figura professionale (dal 4% al 7%). Nel 2017 invece si è assistito ad un aumento considerevole sia del numero di aziende con al proprio interno dei data scientist (dal 31% si è passati al 45%), sia di quelle che riconoscono formalmente questa figura all'interno della propria organizzazione (dal 7% al 16%). La percentuale di

aziende con data scientist è cresciuta, seppur di un punto percentuale, nel 2018 (46)%. Nel 2019 tale percentuale è ulteriormente aumentata al 49%. Il 2020 è il primo anno in cui si assiste ad una stabilizzazione della diffusione del data scientist all'interno delle grandi aziende italiane.

La crescita rilevata in questi anni è il risultato di un costante aumento di consapevolezza delle grandi opportunità celate nei dati: se questi sono affidati a persone con le adeguate competenze, l'adeguata conoscenza del business aziendale e dei processi che lo caratterizzano possono essere colte e trasformate in valore aggiunto. Visti i dati rilevati, per le aziende oggi non avere al proprio interno figure con questo profilo rischia di essere non più solamente una mancata occasione per acquisire un vantaggio competitivo, ma una resistenza anacronistica che può avvantaggiare i competitor che si muovono in anticipo per sfruttare le Big Data analytics.

In sintesi, si può dire che il data scientist presenta cinque diversi "volti":

- 1) Hacker: colui che riesce ad estrarre i dati dalle nuove tecnologie e successivamente li gestisce, li manipola e li predispone all'analisi;
- 2) Scienziato: deve saper impostare esperimenti, progettare gli strumenti da impiegare, raccogliere i dati, analizzarli e descrivere i risultati ottenuti mediante interessanti rappresentazioni dei dati stessi che ne consentano una migliore interpretazione;
- 3) Consulente di fiducia: deve possedere forti competenze di comunicazione e relazione, nonché capacità di formulare decisioni e di capire i processi decisionali;

- 4) Analista quantitativo: deve possedere una conoscenza di modelli e tecniche matematiche, nonché di tool informatici e linguaggi di programmazione mediante cui effettuare analisi;
- 5) Esperto di business: deve possedere una conoscenza precisa di vari aspetti del business aziendale (marketing, finance, produzione, distribuzione, ecc) e del governo di variabili legati al settore in cui l'azienda opera.

Parallelamente all'introduzione di nuove competenze dedicate alla data science, è necessario progettare un nuovo modello di governance delle analytics, per fare in modo che le competenze internalizzate dall'azienda possano essere messe a disposizione delle diverse linee di business.

Le aziende senza un data scientist al proprio interno mostrano un approccio *tradizionale*, orientato a un'analisi dei dati non particolarmente complessa. Le aziende che invece hanno al proprio interno un data scientist si dividono tra quelle che hanno una struttura più *matura*, caratterizzata dalla presenza di più data scientists, ognuno specializzato su un determinato gruppo di competenze, e quelle che invece vedono al proprio interno una presenza più ristretta, limitata ad una figura (una struttura c.d. "*Data Science Enabled*", proprio a sottolineare come l'azienda sia in una fase di sviluppo verso un approccio più diffuso alla data science).

La tendenza registrata negli ultimi anni conferma la nuova consapevolezza delle imprese dei vantaggi dell'integrare al loro interno un team di data scientist: nel 2020, la percentuale di aziende con un modello organizzativo maturo per gestire la data science passa dal 50% del 2019 al 69%, mentre si riduce il numero di aziende che si trovano in un contesto Data Science Enabled (dal 30% del 2019 al 15% del 2020).

1.3 La data economy

Gli algoritmi e l'intelligenza artificiale stanno già trasformando moltissimi ambiti, dall'assistenza ospedaliera alla cura sanitaria a distanza, l'istruzione, i trasporti, l'informazione giornalistica e la fruizione di contenuti multimediali, la cybersecurity, controllo del traffico stradale e previsione del tempo, domotica, Internet delle cose e così via. Possiamo citare solo alcuni esempi (Delmastro e Nicita, 2019) per far comprendere al meglio la portata di questi cambiamenti e le infinite opportunità (legate sempre all'insorgenza di nuovi rischi per l'utente e l'azienda da considerare) che l'applicazione dei Big Data ha comportato.

Nel settore comunicazioni, possiamo riassumere tre passaggi fondamentali rivoluzionati dai Big Data: generazione, traffico e scambio di dati. È possibile oggi scambiare dati con tempi bassissimi di latenza e costi molto ridotti (tra l'altro indiretti e non legati al passaggio di informazioni), passando da "uno-molti" a "molti-molti", alzando l'attenzione del business a ogni forma di comunicazione pubblica e privata. Ecco come come nascono i Big Five (Apple, Microsoft, Google, Amazon, Facebook) e i loro fatturati da impallidire.

Nel settore sanitario, cambia totalmente il modo in cui vengono identificate, trattate e prevenute le malattie, migliorando soprattutto la vita a chi ha diverse o ridotte abilità. Una sanità più attiva, meno costosa e invasiva per la collettività, che permette di assistere a distanza, e ridurre addirittura il numero di interventi grazie a diagnosi precoci o strumenti non invasivi per il paziente.

Nel settore bancario e finanziario, l'elaborazione dei Big Data ha permesso di individuare le transazioni fraudolente, di sviluppare analisi delle tendenze commerciali delle

imprese e di offrire incentivi personalizzati all'uso di carte di credito in funzione dei profili di reddito e di spesa dei clienti. La personalizzazione del rapporto tra istituto di credito e cliente è un elemento decisivo per mantenere o attrarre clientela, tutto tramite registro, archivio, e scambio di dati.

E potremmo citare ancora tantissimi ulteriori esempi tecnici che hanno rivoluzionato altri settori, come la grande distribuzione, l'energia, la logistica e il settore agricolo.

Tutto quanto sopra esemplificato ha in comune quattro elementi: la raccolta di dati rilevati dai comportamenti di imprese e utenti, il loro trattamento da parte di appositi algoritmi, l'elaborazione dei modelli predittivi, la valorizzazione economica dei dati. Questi quattro elementi sono gli ingredienti di un nuovo modello di organizzazione capitalistica dei mercati. Nel passaggio storico che viviamo, questi modelli organizzativi stanno trasformando i mercati tradizionali: si registra, infatti, l'emersione di pochi grandi giganti digitali, i cosiddetti "Big Tech". La preoccupazione circa la concentrazione di dati, algoritmi, probabilità e innovazione nelle mani di pochi colossi a livello globale è oggetto in tutto il mondo di dibattito e discussione da parte di parlamenti, governi, autorità garanti preposte alla tutela della concorrenza, della regolazione dei mercati digitali, privacy e cybersecurity. Sono intrecciate questioni di competitività e affermazione di imprese caratterizzate da dimensioni e potere mai visti in passato, con temi inerenti sia la cybersecurity (connettività in ambienti strategici per paese) sia per manipolazione informativa sul Web per influenzare le campagne elettorali e le scelte dei cittadini. Non sono solo i rischi e i pericoli derivanti dal potere economico di questi colossi a preoccupare le autorità, ma soprattutto l'efficacia delle attuali normative antitrust e di regolazione del mercato del capitalismo digitale, nonché la garanzia dell'equilibrio tra

libertà d'impresa e tutela della concorrenza da una parte, e libertà d'espressione e tutela del pluralismo online dall'altro, con prioritario interesse al cittadino consumatore. Qual è l'equilibrio accettabile tra le opposte esigenze?

1.4 La profilazione del consumatore: tra libertà, canalizzazione della scelta e tutela della privacy

Fino ad ora abbiamo tracciato una linea che sembra avere solo riscontri positivi e margini di miglioramento infiniti per l'economia dei Big Data e di tutte le aziende che partecipano al business. Quello che non sempre viene rilevato e percepito dagli utenti è il rischio legato alla privacy e al trattamento dei dati personali. Avere a disposizione tante dettagliate informazioni è un grosso vantaggio per le aziende ma altrettanto una grossa responsabilità.

Ecco che in qualità di utenti possiamo essere influenzati da alcuni avvenimenti dirottati in campo economico o politico, indirizzati verso un determinato acquisto, ricevere solo pubblicità di un determinato prodotto o ancora più "grave", qualcosa che riguardi i dati sensibili della persona, come l'orientamento religioso, sessuale o politico.

Ma come è possibile (Iandiorio, 2019) che un'azienda o una qualsiasi organizzazione riesca ad arrivare ad avere informazioni così dettagliate sulla vita di una persona solo da un indirizzo IP? Questo avviene quando chi ha questi dati riesce ad incrociarli con altre informazioni provenienti da diverse fonti, come l'attività sui social media, gli acquisti online, le partecipazioni agli eventi, le registrazioni in un luogo con il GPS, ecc.

La trasparenza da parte del soggetto che acquisisce i dati non è equiparata ad un utente consapevole di quali informazioni stia esso trasmettendo, poiché l'utente è sia fruitore che produttore di dati: è necessario sviluppare al meglio un sistema che si occupi della

protezione dei dati sensibili e della tutela della privacy. È plausibile che alcuni prodotti e alcuni servizi siano orientati a persone di un determinato orientamento sessuale o religioso, o che l'accesso ad alcuni servizi finanziari sia precluso ad alcuni privati sotto una determinata fascia di reddito, e tante altre situazioni che favoriscono l'esclusione e la segmentazione dell'uomo non per motivi di interesse reciproco ma prettamente di informazioni acquisite dai dati sensibili. Dati sensibili che potrebbero a volte peccare di mala interpretazione, per motivi tecnici derivanti sia dagli algoritmi che da semplici deduzioni umane non corrispondenti alla realtà soggettiva della situazione. I numeri e i dati non mentono dal punto di vista oggettivo, sia in percentuale che in termini assoluti, ma ci sono alcune condizioni che neanche il miglior software potrebbe recepire nello spiegare un determinato punteggio (influenza dei media, esperienze dirette, manomissioni volontarie, ecc).

Per cercare di ovviare a queste situazioni delicate, il legislatore ha compiuto un enorme passo in avanti nel 2016, quando l'Autorità Garante ha elaborato a livello normativo il Regolamento sui dati personale 2016/679 (GDPR) che, tra le novità principali, ammette l'anonimizzazione del dato come modalità di individuazione del soggetto, e che l'acquisizione o la diffusione degli stessi per fini interni o a terze parti possa avvenire solo tramite consenso espresso dell'utente, che troppe volte però sottovaluta la potenzialità di questi atti.

Non è stato comunque disciplinato il rapporto tra la titolarità del dato e le operazioni riguardanti estrazione e trattamento, approvato solo per fine statistico. Il legislatore separa l'uso per fine statistico ed analitico, che viene approvato solo per "scopi legittimi ed espliciti al loro trattamento, solo se compatibili con la norma". Ma proprio dalla

statistica, e dal successivo cross sources, vengono poste le basi per la canalizzazione della scelta dell'utente: il consumatore viene profilato e canalizzato in una bacheca piena di oggetti e sconti in prima fila, "solo per te" o "solo per oggi", nelle homepage dei giornali online che pubblicano articoli di un dato argomento di interesse, tramite i social media ci arrivano ADS e brevi spot riguardanti prodotti di nostro interesse.

CAPITOLO 2: BIG DATA MARKETING

2.1 Data-driven marketing: l'esigenza di un approccio basato sui dati

Nell'attuale contesto competitivo dominato dai Big Data, estremamente complesso perché mutevole di secondo in secondo, le imprese sono costrette a adottare un approccio al marketing c.d. "data-driven" ("guidato dai dati"), perché ai managers non basta più affidarsi alle proprie conoscenze e alle proprie esperienze, ma necessitano di un'analisi approfondita ed esaustiva dei dati che li circondano. Il dato viene considerato una risorsa fondamentale, indispensabile al fine di prendere decisioni che siano il più accurate e rispondenti alle esigenze aziendali possibile. I dati a disposizione dei decision-maker aziendali sono tanti e tendenzialmente in aumento, per via del continuo sviluppo tecnologico del digitale e del crescente utilizzo di un approccio di marketing "omnicanales", cioè che richiede la completa integrazione di canali fisici e virtuali. Ad aver contribuito alla formazione di questa gigantesca mole di dati sono stati i nuovi strumenti tecnologici quali smartphone, smart device, social media (ecc.), che permettono di condividere, comunicare ed elaborare in tempo reale informazioni di vario tipo.

Dal punto di vista più strettamente tecnologico-informatico, un'impresa che voglia essere competitiva deve dotarsi di:

- ERP (Enterprise Resource Planning): software che supportano la gestione dell'azienda attraverso un'architettura che consente di automatizzare e integrare quantomeno la maggior parte dei suoi processi aziendali;
- CRM (Customer Relationship Management): software CRM che contiene tutte le informazioni relative ai clienti (numeri di telefono, indirizzi, comunicazioni), le quali

vengono successivamente utilizzate per la personalizzazione, la segmentazione, le campagne e altre attività relative alle vendite, al marketing e al servizio clienti;

- SCM (Supply Chain Management): insieme di metodologie gestionali e soluzioni software che consentono di gestire in modo efficiente l'intera catena di distribuzione, monitorando step by step i processi di approvvigionamento, le relazioni con i fornitori, la logistica. Si basa soprattutto sulla previsione, programmazione e coordinamento del flusso delle merci, e sulle aspettative del cliente finale;
- Marketing Operations Management: programma che raggruppa l'insieme di processi e applicazioni che forniscono un quadro strutturale per pianificare, gestire ed eseguire sistematicamente operazioni di marketing quali budgeting, pianificazione del marketing e gestione dei contenuti.

Tuttavia, non è corretto ridurre l'adozione di un approccio data-driven da parte dell'impresa solamente al coinvolgimento di tecnologie (hardware, software, ecc), poiché è fondamentale che vi sia in azienda una predisposizione di tutti, managers e dipendenti, a adottarle ed a saperle utilizzare massimizzandone i benefici.

In sostanza, dopo la raccolta e l'analisi dei dati, si classificano gli "insights" ("intuizioni", "pezzi di conoscenza") più rilevanti ottenuti, poi si progettano le decisioni e le strategie sulla base di questi.

Il "data-driven marketing" viene visto dalla Chief Marketing Officer Lisa Arthur nel suo libro guida pubblicato nel 2013⁵, quindi agli albori della questione, come una grossa

⁵ "Big Data Marketing: Engage Your Customers More Effectively and Drive Value".

opportunità per i dirigenti che devono districarsi dal caos generato dalla crescita incontrollata dei Big Data, che devono essere sezionati e utilizzati da persone esperte all'interno delle divisioni aziendali per supportare le decisioni e il conseguente vantaggio competitivo dell'impresa. Altre interpretazioni considerate rilevanti sono date da Semmelroth (2013)⁶, secondo il quale il "Big Data marketing" non è altro che l'uso delle informazioni sui clienti per condurre correttamente gli sforzi di comunicazione di marketing. Jeffrey (2020)⁷, invece, assume una connotazione più profonda, proponendo indicatori per le informazioni tradizionali e per le informazioni non tradizionali (cioè quelle generate dai nuovi canali digitali): consapevolezza del brand, tasso di abbandono, soddisfazione del cliente, take rate, customer lifetime value per informazioni tradizionali; costo per click (CPC), tasso di conversione delle transazioni (TCR), ritorno economico per ogni dollaro speso in advertising (ROAS), passaparola per informazioni non tradizionali.

In "The new Marketing analytics" (2016), gli autori Dirks e Koechlein sostengono che gli obiettivi da perseguire con il Big Data marketing sono: aumentare la customer acquisition e diminuirne i costi; diminuire il tasso di abbandono; aumentare la profittabilità del cliente; migliorare il prodotto o servizio offerto. Nello stesso testo, gli autori sostengono che il marketing non cambi con l'avvento dei Big Data: si tratta sempre di "sviluppare campagne per fare in modo che i potenziali clienti pensino al prodotto dell'azienda in modo differenziato rispetto alla concorrenza e lo scelgano. Poi

⁶ "Data-Driven Marketing for Dummies".

⁷ "Data-Driven Marketing: The 15 Metrics everyone in Marketing should know".

diventa una questione legata a quanto si è in grado di mantenere questi clienti fedeli nel tempo e, se possibile, aumentare il volume di affari con ognuno di loro”.

Quindi, tenendo conto di tutto quanto detto prima, si può definire il Big Data marketing come “il processo di raccolta, analisi ed esecuzione degli insight dedotti dai Big Data per favorire le relazioni con i clienti, il miglioramento dei risultati di marketing e la misurazione dell’affidabilità interna dell’azienda”. Dunque un processo interno dell’azienda che mira, tramite la sequenza “assetto informatico - estrapolazione dei dati - decisioni”, e grazie ai risultati delle “grandi informazioni” generate, facilitate ed elaborate, a raggiungere una migliore comprensione dei gusti, dei desideri e delle esigenze del consumatore.

In sintesi, si può affermare dunque che, affinché il Big Data marketing possa essere applicato in una azienda, sono essenziali sei funzionalità da sviluppare continuamente (Boston Consulting Group, 2014)⁸:

- Opportunità: costruire una cultura dell’innovazione e della sperimentazione;
- Fiducia: stabilire la fiducia tra i consumatori, per consentire un uso più ampio delle loro informazioni;
- Piattaforma: sistemi informativi flessibili, scalabili ed efficienti;
- Organizzazione: sviluppo di capacità per l’implementazione e la relativa leva delle applicazioni informative;
- Partecipazione: identificazione di partner strategici che possano aiutare a sbloccare nuove opportunità economiche;

⁸ “Enabling Big Data: Building the Capabilities that really matter”.

- Relazioni: creare una cultura aperta di supporto tra i partner ed essere aperti alla condivisione di informazioni.

2.2 Big Data analytics: un nuovo strumento per il supporto alle decisioni

Con l'espressione "Big Data analytics" si intende l'insieme di processi che permettono di compiere la trasformazione dei dati in conoscenza immediatamente utilizzabile in azioni favorevoli all'impresa. Il loro scopo fondamentale è quello di esplorare i dati e identificare modelli e relazioni impercettibili a prima lettura per ottenere approfondimenti avanzati sugli utenti mediante tecnologie correlate per il miglioramento (valorizzazione dal punto di vista informatico) e la visualizzazione dei dati. Secondo Mark and Douglas (2012), le Big Data analytics si basano su metodi statistici, previsioni, analisi di regressione, interrogazione di database e magazzinaggio di dati.

"Analytics" è un "umbrella term for data analysis applications"⁹ (un termine che copre, come un ombrello aperto, tutte le applicazioni di analisi dei dati), ed è quindi un concetto generale applicabile a tutte le tecniche di analisi quantitativa a supporto delle decisioni.

Per sfruttare al meglio il potenziale dei Big Data, è necessario estrapolarne il significato intrinseco: l'azienda deve capire quale tipo di analytics è il più adatto per il proprio business. Esistono quattro tipologie di Big Data analytics, in ordine crescente di livello di impatto sull'azienda:

⁹ Mandelli, 2017.

- Descriptive analytics (analisi descrittive), le quali racchiudono gli strumenti e le tecniche in grado di descrivere la situazione attuale e passata dei processi aziendali e/o fenomeni di interesse, come ad esempio la riduzione delle vendite in un determinato periodo;
- Diagnostic analytics (analisi diagnostiche), le quali catturano relazioni tra i dati e supportano lo sviluppo di ipotesi;
- Predictive analytics (analisi predittive), le quali permettono di anticipare eventi per fornire maggior controllo sulla loro gestione avvalendosi di tecniche statistiche (regressione, forecasting, modelli predittivi, ecc);
- Prescriptive analytics (analisi prescrittive), le quali suggeriscono soluzioni ottimali a problemi generali e contingenti, in base a parametri dati.

Elisa Landiorio aggiunge poi le “automated analytics” (analisi automatizzate), definendole come “tecniche e tools in grado di implementare automaticamente e autonomamente un’azione o un modello di azioni che viene proposto a seconda del risultato delle analisi”: è un tipo di analisi strettamente legata ai nuovi sistemi di intelligenza artificiale.

L’analisi descrittiva si basa su tecniche di data mining che mirano ad analizzare e tradurre rapidamente e in tempo reale grandi set di dati strutturati grezzi (ad esempio numero di post, menzioni, fan, follower, visualizzazioni di pagina, check-in, ecc) dei clienti, che di per sé non hanno alcun significato, in informazioni preziose per facilitare i processi decisionali: i modelli descrittivi consentono di apprendere dai comportamenti o dalle esigenze passate dei propri clienti e di comprendere l'impatto di tali dati sui comportamenti o sui bisogni futuri.

L'analisi predittiva utilizza, per lo studio dei dati non strutturati dei clienti (a differenza dell'analisi descrittiva) sia storici che recenti, una varietà di tecniche statistiche e matematiche, di data mining e di machine learning, che consentono di fare previsioni sul futuro del business aziendale, sulla domanda futura di un prodotto, sui modelli di acquisto o su altri comportamenti dei clienti.

L'analisi prescrittiva, grazie alla tecnologia, supera i modelli descrittivi e predittivi, formulando soluzioni e suggerimenti che consentono di effettuare scelte e decisioni di business in modo più rapido ed efficace.

Secondo una ricerca del 2017 dell'Osservatorio Big Data Analytics & Business Intelligence del Politecnico di Milano, nelle grandi aziende con più di 250 dipendenti la descriptive analytics è presente nell'89% dei casi, mentre la predictive analytics solo nel 59% (per il 30% riguarda solo alcuni ambiti e un 28% addirittura è solo un progetto pilota); il restante 41% delle aziende non ha ancora soluzioni di questo tipo, tuttavia il 37% ne sta valutando l'adozione nei prossimi 12 mesi. Sono ancora molto indietro le prescriptive e automated analytics, presenti rispettivamente nel 23% e nel 10% delle aziende e solo a livello di progetto pilota.

Per meglio comprendere l'impatto dei Big Data e delle Big Data analytics sulle varie attività di marketing, affinché le aziende possano sfruttarne meglio i benefici, si fa riferimento alla c.d. "teoria basata sulle risorse" (in inglese "RBT", ovvero "resource-based theory"), secondo la quale tre tipi di risorse quali capitale fisico, capitale umano e capitale organizzativo moderano tre fasi di un processo sequenziale:

- la raccolta e l'archiviazione delle prove dell'attività dei consumatori come Big Data;
- l'estrazione di informazioni sui consumatori dai Big Data;

- l'utilizzo di tali informazioni per migliorare le capacità dinamiche/adattive.

Le risorse di capitale fisico includono software o piattaforme per raccogliere, archiviare o analizzare i Big Data. Le risorse di capitale umano includono le conoscenze dei data scientist nell'acquisire, gestire ed estrarre informazioni dalle attività dei consumatori. Le risorse di capitale organizzativo includono una struttura organizzativa che consenta all'azienda di trasformare le intuizioni dei Big Data in azioni, in quanto potrebbe aver bisogno di modificare l'organizzazione e i processi aziendali.

Sebbene i Big Data siano oggi considerati una nuova forma di capitale, molte aziende non riescono a sfruttarne i benefici, in quanto non riescono ad allocare adeguate risorse dei tre tipi di capitale. Più i dati diventano numerosi e complessi, più le limitate capacità mentali degli esseri umani fanno difficoltà nel decifrare e interpretare un ambiente sconosciuto. Per questo motivo si è verificato un ribaltamento del metodo scientifico, nel senso che dall'adattamento dei dati alle teorie del mercato si è passati all'utilizzo dei dati per inquadrare le teorie: si fa meno affidamento sulle conoscenze esistenti e invece si fa maggiore attenzione su ciò che è sconosciuto. In sostanza, si passa da una visione basata sulla conoscenza ad una visione basata sull'ignoranza.

La resource-based theory offre una preziosa spiegazione dell'impatto dei Big Data sul marketing: le risorse (tangibili e non) di un'impresa facilitano le sue prestazioni e il vantaggio competitivo quando sono preziose (cioè quando migliorano i profitti di un'azienda o quando generano un valore per i clienti che i concorrenti non possono raggiungere), rare (non abbondanti), imperfettamente imitabili (non facilmente copiabili) e sfruttabili (in un modo in cui altri non possono sfruttarle) dall'impresa.

La capacità di un'impresa di rispondere ai continui cambiamenti dell'ambiente ipercompetitivo esterno (c.d. capacità "dinamica") impone di incorporare abilità e conoscenze interne per aggiornare e riconfigurare le risorse esistenti e creare nuovo valore: un'impresa che utilizza nuove informazioni sui consumatori estratte dai Big Data per comprendere le esigenze dei consumatori non soddisfatte, infatti, migliora la capacità dinamica¹⁰. Le imprese dovrebbero poi anche assumere un atteggiamento proattivo nel rispondere ai cambiamenti dell'ambiente esterno, senza variare la struttura organizzativa, ma catturando segnali dall'attività dei consumatori, per estrarne intuizioni nascoste al fine di prevedere poi le loro tendenze e quelle del mercato (c.d. capacità "adattiva")¹¹.

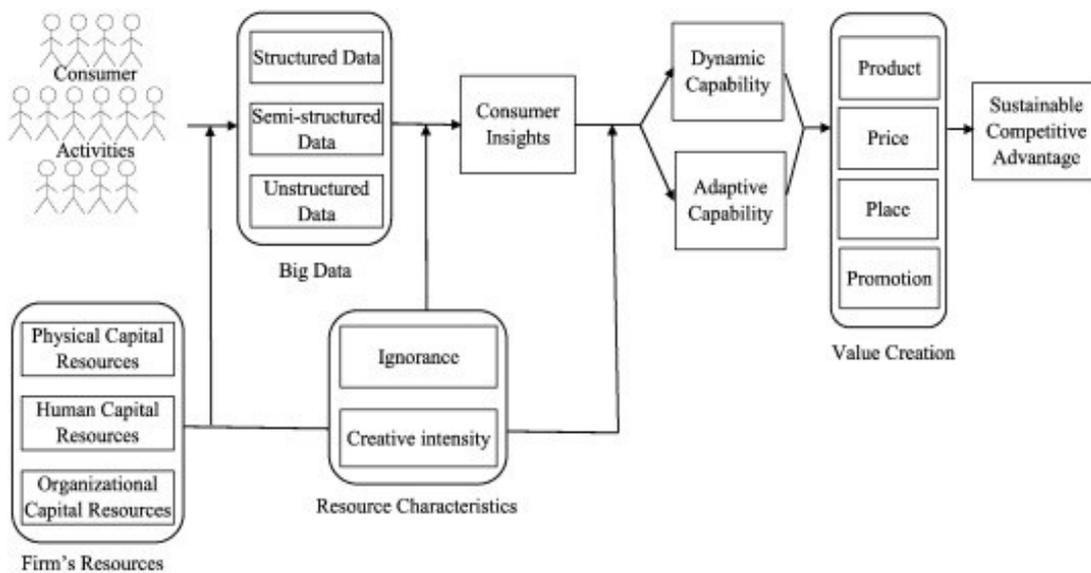
Uno dei limiti della resource-based theory riguarda l'origine delle risorse e delle capacità: la soluzione è vista nel concetto di "ignoranza" come risorsa unica richiesta dalle aziende per ottenere un vantaggio competitivo sostenibile utilizzando i Big Data. A sua volta, infatti, l'ignoranza consente la conoscenza, perché realizzare ciò che non si conosce è fondamentale per la scoperta di nuove conoscenze ("capire ciò che non sappiamo, è importante quanto comprendere ciò che sappiamo"). L'ignoranza stimola la creatività all'interno di un'impresa, mentre una conoscenza perfetta può ostacolare le attività creative. In tale contesto, quindi, una visione basata sull'ignoranza piuttosto che sulla conoscenza consente di facilitare la scoperta di intuizioni nascoste nei Big Data.

¹⁰ La compagnia aerea Southwest Airlines registra le conversazioni tra il personale di servizio e i passeggeri per estrarre informazioni utili mediante software di analisi vocale.

¹¹ La società di distribuzione Target è in grado di stimare se un acquirente donna è incinta; di conseguenza, utilizza i dati per influenzare i loro acquisti in prodotti per bambini prima dei concorrenti.

La figura 1 riassume il processo di creazione di un vantaggio competitivo sostenibile da parte di un'impresa: tramite la giusta allocazione delle risorse, l'impresa riesce a raccogliere dati dalle attività dei consumatori, che successivamente trasforma in informazioni utili mediante una visione basata sull'ignoranza, che stimola la creatività; infine, grazie alla sua capacità dinamica e a quella adattiva, l'impresa riesce a creare valore e, di conseguenza, un vantaggio competitivo sostenibile.

Figura 1: Il processo di creazione di un vantaggio competitivo sostenibile



Fonte: SUNIL EREVELLES, NOBUYUKI FUKAWA, LINDA SWAYNE - Big Data consumer analytics and the transformation of marketing, Journal of Business Research, Volume 69, Issue 2, 2016, Pages 897-904.

Nell'articolo "Big Data and firm marketing performance: Findings from knowledge-based view", al fine di identificare i modi in cui le Big Data analytics possono essere utilizzate per comprendere e migliorare le prestazioni di marketing di un'azienda, gli autori hanno intervistato 10 individui di vertice in varie aziende in diversi settori in India. I risultati che indicano molti di loro presentavano un deficit cognitivo, in quanto non avevano la capacità tecnica di combinare, costruire e riorganizzare le competenze

interne ed esterne essenziali per utilizzare i Big Data, sebbene i principi fondamentali dell'utilizzo dei dati per aggiornare il processo decisionale non variassero in modo significativo. Nonostante gli intervistati riconoscano il potenziale dei Big Data per migliorare la strategia di marketing, generalmente si consideravano mal equipaggiati per gestirli, credendo fosse una competenza delle generazioni più giovani.

Molti manager presentavano un bias (pregiudizio) cognitivo, poiché erano sospettosi dei Big Data ("c'è costantemente il dubbio sull'uso delle informazioni e dei dati [...] e apprezzo che ci sia un aspetto di prudenza"). Tutti hanno espresso la preoccupazione che i dati possano essere interpretati erroneamente ("Non sono sicuro che i boards riescano sempre a catturare i Big Data").

Gli intervistati hanno poi percepito l'onere cognitivo nell'analizzare grandi set di dati (più grandi di quelli che possono competentemente elaborare): sebbene uno di loro abbia notato la "vicinanza ... e ... l'interrelazione" dei modelli di dati, ha anche suggerito che le aziende potrebbero incontrare troppi "Big Data per essere veramente competenti per analizzare con successo".

Alcuni hanno indicato che i Big Data potrebbero non sempre produrre approfondimenti di alta qualità: sebbene gli strumenti digitali abbiano trasformato o migliorato la quantità di dati accessibili, "la pratica decisionale è molto basata [...] su informazioni di migliore qualità, informazioni molto vigorose, che possono oscurare cosa sta realmente accadendo". I dati in eccesso possono ostacolare la capacità degli operatori di marketing di "essere veramente competenti per analizzarli con successo".

I manager hanno espresso preoccupazione per la capacità dei consigli di amministrazione di gestire i Big Data, poiché sono sotto pressione "per trasformare le

loro aziende il più rapidamente possibile". Secondo loro, inoltre, le incongruenze nei Big Data potrebbero distrarre dal processo decisionale strategico, alimentando le preoccupazioni riguardo alle capacità di analisi dei consigli di amministrazione e ai conseguenti impatti negativi sull'azienda: "Riscontriamo regolarmente circostanze in cui ci sono grandi differenze nei dati, poiché la valutazione degli oggetti è confusa, a volte è davvero difficile". I consigli di amministrazione, infatti, tendono a essere composti da uomini di mezza età che generalmente non hanno conoscenze tecniche o familiarità con grandi database e spesso resistono alle nuove tecnologie.

Tutti hanno generalmente riconosciuto l'utilità delle Big Data analytics per esporre le possibilità strategiche indagando su dati tecnici e di mercato esterni in combinazione con competenze e attività interne ("accelera, migliora la competenza umana. Può sviluppare più intuizioni con il personale, collegare spazi tra intuizioni e miglioramenti. Potrebbe identificarsi meglio con le persone. Può sviluppare contatti e soluzioni per problemi aziendali").

2.3 Big data analytics in una prospettiva di marketing mix

Un'interessante panoramica delle Big Data analytics nel marketing è quella proposta nell'articolo "Demystifying Big Data Analytics for Business Intelligence Through the Lens of Marketing Mix": per ciascuna variabile del marketing mix (in aggiunta alle classiche "4P" dell'omonimo modello di marketing mix quali prezzo, prodotto, comunicazione e distribuzione viene aggiunta una quinta "P" di "persona") vengono identificate le fonti dei dati che l'impresa raccoglie, i metodi che utilizza per analizzarli, rielaborarli e tradurli e infine le applicazioni destinate a questi dati.

Figura 2: Il panorama delle Big Data analytics visto attraverso il modello del marketing mix.

	People	Product	Promotion	Price	Place
Data	<ul style="list-style-type: none"> Demographics Social Networks Customer Review Click Stream Survey Data 	<ul style="list-style-type: none"> Product Characteristics Product Category Customer Review Survey Data 	<ul style="list-style-type: none"> Promotional Data Survey Data 	<ul style="list-style-type: none"> Transactional Data Survey Data 	<ul style="list-style-type: none"> Location-based social networks Survey Data
Method	<ul style="list-style-type: none"> Clustering Classification 	<ul style="list-style-type: none"> Association Clustering Topic Modeling 	<ul style="list-style-type: none"> Regression Association Collaborative Filtering 	<ul style="list-style-type: none"> Regression Association 	<ul style="list-style-type: none"> Regression Classification
Application	<ul style="list-style-type: none"> Customer Segmentation Customer Profiling 	<ul style="list-style-type: none"> Product Ontology Product Reputation 	<ul style="list-style-type: none"> Promotional Marketing Analysis Recommender Systems 	<ul style="list-style-type: none"> Pricing Strategy Analysis Competitor Analysis 	<ul style="list-style-type: none"> Location-based Advertising Community Dynamic Analysis

Fonte: SHAOKUN FAN, RAYMOND Y.K. LAU, J. LEON ZHAO - Demystifying Big Data Analytics for Business Intelligence Through the Lens of Marketing Mix (2015) pp. 28-32.

Come si vede dalla figura n.2, il sondaggio (survey) è il metodo più comune per acquisire dati: è definito come “la raccolta di informazioni in modo organizzato e metodico sulle caratteristiche di interesse da alcune o tutte le unità di una popolazione utilizzando concetti, metodi e procedure ben definiti e compilare tali informazioni in un utile modulo riassuntivo”.

Le tecniche di data mining possono aiutare a sviluppare informazioni dai dati per il processo decisionale di marketing estraendo o rilevando modelli o prevedendo il comportamento dei clienti da database di grandi dimensioni: i metodi più comuni di data mining includono l’associazione, la classificazione, il clustering, la regressione, il topic modeling e il collaborative filtering. Ovviamente, vanno selezionati i metodi di data mining appropriati in base alle caratteristiche dei dati e ai problemi aziendali.

PEOPLE

Per una strategia di marketing efficace, è essenziale segmentare¹² i clienti in gruppi specifici che condividano preferenze e interessi simili e rispondano a un segnale di marketing specifico. Le strategie di marketing si basano principalmente sulla conoscenza che un'azienda ha acquisito dei propri clienti (c.d. "customer knowledge", ovvero la "combinazione di esperienza, valore e informazioni approfondite necessarie, create e assorbite durante la transazione e lo scambio tra i clienti e l'impresa"), che viene utilizzata per identificare gli individui a cui è più probabile che il prodotto attiri e venda. Pertanto, è fondamentale per le aziende stabilire un processo di STP ("Segmentazione-Targeting-Posizionamento") forte e coerente. Tuttavia, la segmentazione dei clienti sta diventando sempre più difficile per via dell'enorme volume e dell'immensa varietà dei dati stessi: ora, con il marketing più mirato e personalizzato (c.d. "marketing one-to-one"), non si tratta più di identificare gruppi di clienti simili ma di "profilare" ogni singolo cliente in base ai suoi dati, in modo che i beni e servizi più adatti siano commercializzati all'individuo più appropriato. In pratica, si raccolgono i dati degli utenti da articoli online, post di blog, social network, discussioni di forum, piattaforme di e-commerce (ecc) e si monitora l'evoluzione dei loro interessi al fine di identificare le tendenze chiave (informazioni, domande o commenti) sui propri prodotti.

Il marketing one-to-one enfatizza quindi la personalizzazione dei messaggi e delle offerte di prodotti per un singolo consumatore. La personalizzazione apre nuove

12 Un interessante modello di segmentazione è il c.d. "modello RFM" (Recency – Frequency – Monetary Value): la recency (prossimità nel tempo) è l'intervallo di tempo dall'ultima operazione di acquisto; la frequenza indica quanto spesso il cliente acquista; il valore economico indica quanto denaro il cliente ha lasciato nella cassa aziendale.

opportunità di business, in particolare i motori di raccomandazione del prodotto (PRE), sistemi di filtraggio dei dati che classificano gli oggetti in base alla loro rilevanza per l'utente, creando uno schema di preferenze per prevedere il punteggio che l'utente assegnerebbe a ciascun oggetto e infine raccomandando l'oggetto con il punteggio più alto.

Sempre più aziende, poi, si affidano alle Big Data analytics per migliorare l'esperienza del consumatore (c.d. "customer experience", ovvero "le percezioni del cliente - sia cosce che inconscie - della sua relazione con il marchio risultante da tutte le loro interazioni durante il ciclo di vita del cliente stesso") come mezzo per aumentarne la soddisfazione e il coinvolgimento: le analytics consentono, infatti, anche di comprendere i sentimenti dei clienti e creare con loro delle connessioni emotive, basandosi su sondaggi formali, trascrizioni delle chiamate, commenti sui social network e qualsiasi altro tipo di scambio tra un marchio e un cliente. A tal fine, dei sistemi di messaggistica basati sui Big Data, chiamati "chatbot", sono stati implementati su siti Web, forum e piattaforme social, per interagire direttamente in tempo reale con i propri clienti fornendo loro informazioni o assistenza pertinenti. Oltre che facilitare il dialogo con i clienti, tali chatbot permettono di ridurre i costi assistenza.

PRODUCT

Per poter recuperare, e successivamente gestire, la reputazione di un prodotto a seguito dell'analisi dei dati dei sondaggi, sono stati proposti diversi metodi: in particolare, Di et

al.¹³ ha proposto un metodo che dal Web non solo estrae dati testuali ma considera anche le immagini del prodotto ivi pubblicate.

Per eseguire un'analisi automatica dei commenti testuali e delle immagini pubblicate sul Web è essenziale però sviluppare prima una ricca “ontologia” del prodotto, ovvero una rappresentazione computerizzata (computer-based) delle informazioni riguardo quel determinato prodotto. Recentemente, è stato sviluppato un metodo¹⁴ automatizzato consistente nella modellazione di argomenti latenti per costruire ontologie di prodotto basate su descrizioni testuali estratte dai social media; tuttavia, è necessario sviluppare continuamente nuovi metodi computazionali di estrazione per far fronte ai problemi di volume, velocità e varietà dei dati provenienti dai social media.

PROMOTION

I dati promozionali di solito includono informazioni sul tipo di promozione (riduzione del prezzo o coupon), sul tempo di promozione e sul record di acquisto durante il periodo promozionale.

L'analisi di marketing promozionale può includere anche altri fattori come il prezzo e il luogo: ad esempio le aziende possono utilizzare le informazioni sulla posizione dei clienti per migliorare la propria strategia di promozione e selezionare clienti mirati.

La posizione è una delle soluzioni più importanti per soddisfare le esigenze dei consumatori ed è una preziosa fonte di informazioni di marketing personalizzate: con la pubblicità basata sulla posizione, i clienti possono ricevere annunci pubblicitari

¹³ Is a picture really worth a thousand words?: on the role of images in e-commerce. Proceedings of the 7th ACM International Conference on Web Search and Data Mining, ACM (2014), pp. 633-642.

¹⁴ Social analytics: learning fuzzy product ontologies for aspect-oriented sentiment analysis. Decis. Support Syst., 69 (2014), pp. 80-94.

tempestivi o consigli sui prodotti in base alla loro posizione attuale o alla posizione futura prevista.

PRICE

Diverse strategie di prezzo sono preferite in diverse situazioni di marketing: ad esempio, la crescita dell'e-commerce ha reso disponibili le informazioni sui prezzi direttamente sui siti Web e i ricercatori hanno iniziato a utilizzare i dati di registro per studiare, tramite regressioni, la strategia dei prezzi nei siti di e-commerce.

Un'applicazione di analisi della concorrenza automatizzata tramite metodi di associazione non si limita a identificare i potenziali concorrenti di un'azienda, ma scopre efficacemente anche i prodotti potenzialmente competitivi nei contesti di prodotto.

PLACE

Oggi, i Big Data sono la risorsa e al contempo lo strumento più importante per la gestione della catena di distribuzione e l'ottimizzazione dei costi della logistica: la loro irruzione trasforma il modus operandi delle aziende in modo considerevole, obbligandole a ripensare completamente le dinamiche di magazzino.

Molte aziende oggi hanno a disposizione diverse nuove tecnologie che cambiano radicalmente il sistema della catena di distribuzione, abbattendo i costi della logistica: robot, droni, veicoli a guida autonoma, stampanti 3D (ecc).

I Big Data permettono di analizzare l'efficacia delle proprie strategie di material handling, di studiare l'efficienza dei flussi di materiali (inbound e outbound), di migliorare le previsioni di vendita e di offrire ai clienti un servizio di consegna

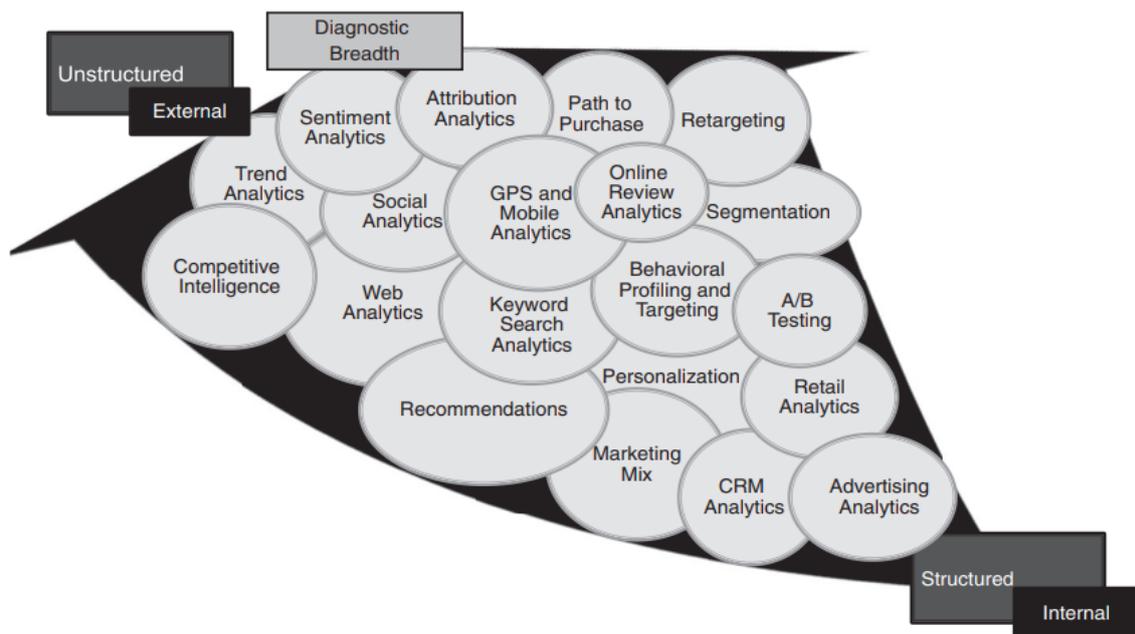
personalizzato: tutto ciò si traduce a livello pratico in una migliore distribuzione ed allocazione delle merci all'interno e all'esterno di un magazzino.

Il machine learning sta rivoluzionando la logistica: gli attuali WMS (warehouse management system) sono capaci di analizzare lo storico dello stock, elaborare la migliore mappatura di magazzino e analizzare i percorsi dei camion per ottimizzare i tragitti.

2.4 Big Data analytics nel marketing: applicazioni

Il campo di applicazione delle Big Data analytics nel marketing è ben rappresentato nella figura 3. La freccia mostra come l'analisi di marketing dei Big Data crei una crescente ampiezza diagnostica delle informazioni, spesso vantaggiosa per supportare obiettivi a lungo termine, in funzione dell'utilizzo di dati interni (soprattutto strutturati) e dati esterni (soprattutto non strutturati).

Figura 3: L'ampiezza diagnostica delle Big Data analytics nel marketing.



Fonte: WEDEL M, KANNAN P.K. - Marketing Analytics for Data-Rich Environments (2016) pp. 97-121.

SOCIAL MEDIA ANALYTICS

La social media analytics consiste nella raccolta, nell'integrazione e nell'analisi dei dati generati dai social network, sia quelli più famosi come Facebook, LinkedIn, YouTube, Twitter (ecc), ma anche di tutti i social media del Web (ad esempio giornali online, forum e blog): attraverso tali dati si possono analizzare i comportamenti dei propri visitatori (che siano clienti attuali, clienti potenziali o semplici seguaci) al fine di ottenere un targeting molto preciso. I dati ricavati possono essere di varia natura: testuali (commenti, post, ecc); dati relativi al network analizzato (amicizie di Facebook, sistema di following di Twitter o Instagram, ecc); azioni (condivisioni, reazioni, "mi piace, ecc).

Esistono due tipi di social media analytics:

- Analisi qualitative, le quali esplorano gli aspetti più intangibili della presenza online delle persone, quali il posizionamento, il tono di voce, la qualità dell'interazione con follower e fan, la tipologia e la qualità dei contenuti utilizzati, fotografie, video, testi. Nell'analisi su tematiche specifiche, consentono di comprendere il *sentiment*, ovvero le opinioni e le tendenze prevalenti rispetto a un tema, a un personaggio, a un brand;
- Analisi quantitative, le quali valutano, all'opposto, le performance di tutte le dimensioni misurabili della presenza social degli individui, attraverso un'analisi quantitativa delle conversazioni svolta mediante l'applicazione di algoritmi di Natural Language Processing (NLP), text analysis e analisi biometriche.

Grazie all'analisi dei social media, in sostanza, è possibile: ricavare informazioni demografiche degli utenti (sesso, età, residenza, ecc) e monitorarne il comportamento generale; individuare le keyword più ricercate e usate, anche relativamente a un

determinato articolo e servizio in vendita, o su determinati post, link, ecc; verificare orari, giorni e luoghi in cui il social viene visitato maggiormente dagli utenti.

Le attività di social media analytics si suddividono generalmente in tre fasi.

L'identificazione dei dati è il primo essenziale passaggio, perciò vanno individuati fin da subito gli attributi principali da considerare: il tipo di contenuto (testo, audio, video, immagine, ecc); la fonte da cui quel dato proviene (social media, sito Web, ecc); il proprietario del dato (utente pubblico o privato). La fase successiva è l'analisi dei dati mediante i programmi preinstallati nei social media, i quali sviluppano un modello di dati che organizza i contenuti e dà una panoramica di come essi interagiscono fra di loro (ad esempio, il Wordcloud). L'ultima fase è quella dell'interpretazione delle informazioni ricavate dall'analisi: è consigliato organizzarle in un grafico, tenendo bene presente il target (il pubblico a cui ci si rivolge), poiché esistono molti tipi di pubblico e, di conseguenza, diversi tipi di interazioni con essi.

RETAIL ANALYTICS

La possibilità di studiare il comportamento dei consumatori all'interno e all'esterno dello store permette ai responsabili dei punti vendita di intercettare tutte le dinamiche di interazione sia off line che on line dei clienti, attuali e potenziali. La progressiva adozione di una serie di tecnologie rappresenta un patrimonio informativo preziosissimo per i retailer: i beacon (piccoli dispositivi che "parlano" con le persone che transitano nelle vicinanze del punto vendita tramite gli smartphone), i sensori biometrici (sistemi informatici capaci di identificare gli individui a partire dalle loro caratteristiche fisiologiche e comportamentali), le soluzioni di digital signage in store (schermi elettronici con cui è possibile interagire), i siti di e-commerce, le app a

supporto dei programmi di fidelizzazione, i chatbot, i call center, la gestione degli accessi e i dati di scontrinaggio.

Con “retail analytics” si intende l’utilizzo di strumenti analitici predittivi per eseguire analisi delle tendenze, degli eventi e delle performance nel settore della vendita al dettaglio, con l’obiettivo di migliorare l’esperienza del cliente lungo tutto il suo percorso d’acquisto (“path to purchase”), aumentare le vendite e ottimizzare le operazioni.

Ma con quali tipi di dato? In primo luogo, i dati del punto vendita: non solo i dati relativi al venduto, ma anche i margini di profitto, il conteggio dei clienti e le tendenze di vendita. Un’ulteriore fonte di dati sono i sondaggi online, per comprendere le preferenze e l’orientamento dei propri clienti, e le ricerche di mercato, per ottenere informazioni utili su clienti, concorrenti e mercato di destinazione. I social media sono una fonte affidabile di informazioni, purché si agisca sugli stessi canali frequentati dal proprio pubblico di riferimento. In ultima istanza, sono utili anche i dati del traffico pedonale e quelli del traffico Web.

Riassumendo, mediante questo tipo di analisi un retailer può:

- 1) Personalizzare l’esperienza d’acquisto dei clienti con offerte mirate e altamente personalizzate;
- 2) Gestire più efficacemente l’inventario e la supply chain per aumentare l’efficienza operativa;
- 3) Segmentare i clienti per affinità e attributi comuni, comprendendone il profilo e la cronologia su tutti i canali e monitorandone le attività;

- 4) Creare modelli predittivi per collegare comportamenti e dati demografici passati e valutare la propensione d'acquisto futura;
- 5) Implementare strategie di marketing mirate ed efficaci (data-driven), riducendo il numero di decisioni basate su istinti o congetture.

WEB ANALYTICS

La definizione ufficiale di Web analytics, proposta dall'associazione internazionale del settore (la Digital Analytics Association), è "la misurazione, la raccolta, l'analisi e il reporting dei dati Internet allo scopo di capire e ottimizzare l'utilizzo del Web". In altri termini, si può definire la Web analytics come l'insieme di tecnologie e metodologie di analisi del Web che permettono di conoscere più miratamente e dettagliatamente gli utenti del proprio sito, al fine di rafforzare la presenza del brand sul Web, valorizzando l'esperienza degli utenti, fidelizzando chi già conosce l'azienda e aumentando le possibilità di acquisire nuovi clienti.

La Web analytics permette di individuare un'estesa e specifica mole di informazioni:

- informazioni quantitative (ad esempio quanti visitatori o quanti click su banner promozionali interni);
- informazioni qualitative (ad esempio le pagine che l'utente ha visitato o le parole chiave che ha utilizzato nella ricerca);
- la provenienza e la geolocalizzazione (da dove arrivano e dove sono fisicamente);
- la distribuzione nel tempo (in quale momento e per quanto tempo);
- il livello di user experience (le azioni di conversione o abbandono da parte dell'utente).

La maggior parte dei processi di Web analytics prevede quattro fasi essenziali:

- Raccolta dei dati attraverso software specifici (ad esempio Google Analytics):
numero dei visitatori al sito, pagine viste, canali di provenienza del traffico, ecc;
- Elaborazione dei dati in informazioni (in particolare in metriche) utili a pianificare investimenti per migliorare il sito e le azioni di marketing online;
- Sviluppo dei KPI (“key performance indicators”, ovvero indicatori chiave di performance), un insieme di misure quantificabili che un'azienda utilizza per determinare il progresso nel conseguimento dei suoi obiettivi strategici e operativi, e anche per confrontare le sue prestazioni in relazione ad altre aziende all'interno del suo settore;
- Formulazione della strategia online.

RETARGETING

Anche se una strategia STP (Segmentazione-Targeting-Posizionamento) viene implementata con successo, alcuni consumatori mirati potrebbero non essere convertiti in clienti: in questi casi, le tecnologie possono aiutare a riorientare quei consumatori mediante un processo di marketing chiamato “re- targeting”, che generalmente comporta il targeting di chi ha visitato un sito Web ma non ha effettuato un acquisto. Questa pratica è più comunemente applicata nel campo dell’e-commerce: infatti, un potenziale cliente che ha consultato uno o più cataloghi di prodotti senza convalidare l'acquisto è successivamente esposto a pubblicità personalizzata che evidenzia il prodotto inizialmente consultato.

Tale processo è realizzato attraverso l'inserimento di cookie di tracciamento forniti dal sistema di retargeting: vengono offerti incentivi mirati per rilanciare l'interesse iniziale

di un consumatore per un prodotto sulla base di sistemi di raccomandazione e di un'analisi in tempo reale del comportamento del cliente e della cronologia online.

I programmi di retargeting *basico* coinvolgono il consumatore durante tutto il percorso di acquisto, mostrandogli annunci del prodotto che ha visualizzato di recente. I

programmi di retargeting *avanzato*, invece, presentano annunci dinamici e personalizzati, basati sugli articoli preferiti dai clienti (tra questi vi possono essere articoli che un consumatore potrebbe non aver ancora visto, ma che hanno elevate probabilità di essere acquistati).

Il retargeting *statico* prevede la creazione di un gruppo fisso di annunci, presentati al consumatore a seconda delle pagine che ha visitato sul proprio sito. Il retargeting *dinamico* crea annunci personalizzati per ciascun consumatore: richiede una tecnologia più sofisticata per analizzare i comportamenti dei consumatori, per realizzare l'annuncio in base alle preferenze individuali e per stabilire quando presentare l'annuncio e massimizzare così le probabilità di conversione.

Il costo reale di una campagna di retargeting dipende dal costo di collocazione di ciascun annuncio e, naturalmente, dal budget stabilito. Normalmente, le campagne di retargeting vengono offerte in base a modelli di CPM (costo per mille impression) o di CPC (costo per clic). Il CPM è la scelta migliore se l'obiettivo è la brand awareness, dal momento che si paga per le impression ma non si ha la garanzia per ogni clic. CPC significa che si paga solo quando qualcuno fa clic, poiché è probabile che chi clicca sia molto interessato all'offerta.

Tra le altre applicazioni delle Big Data analytics nel marketing si possono segnalare le seguenti:

- GPS and mobile analytics: la tecnologia GPS offre dati di enorme valore all'azienda che può sfruttare questo tipo di applicazione. Consiste principalmente nel geo-targetizzare i clienti con offerte promozionali basate su contesti situazionali, dato che in maniera oggettiva e chiara identifica tutti gli spostamenti del device su cui è stato localizzato per la prima volta il cliente, tra cui casa, lavoro, divertimento, sport, ecc. L'azienda può dunque generare un profilo cliente basato sui suoi spostamenti, che le permette di monitorare i suoi rapporti con la concorrenza, la provenienza, i tempi di permanenza nel suo negozio o negli altri, ecc;
- Online review analytics: come piccolo ma importante tassello delle Web analytics, questa applicazione nel corso degli ultimi anni è diventata sempre più rilevante nell'influenzare i consumatori, soprattutto per quanto riguarda alcuni settori che hanno visto la nascita di aziende specializzate nel raccogliere recensioni (TripAdvisor, Booking, Jobreferences, ecc). Dal punto di vista aziendale, è rilevante per l'impresa avere una solida base di recensioni positive, che valorizzino il brand e l'immagine della stessa, così come può destare preoccupazione il ricevere valutazioni negative riguardo i propri prodotti e/o servizi (queste ultime possono rappresentare anche un'opportunità di rilancio). Predisporre applicazioni sia interne che esterne per "regolarizzare" la sezione recensioni diventa quindi uno strumento valutabile per rendere trasparente l'immagine aziendale;
- Trend analytics: assieme alla competitive intelligence, la trend analytics aiuta le aziende a identificare i cambiamenti nell'ambiente e impostare le difese per mantenere la quota di mercato. Identificabile come ramo della text analytics, questa applicazione studia quali sono gli hashtag del momento (gli aggregatori

tematici che canalizzano le ricerche dei clienti su determinati aspetti o informazioni). Spesso non è così facile prevedere e applicare una strategia basata esclusivamente sul trend ma, se monitorato costantemente e soprattutto se riferito agli eventi passati, mette sicuramente a conoscenza l'impresa di ciò che il mercato vuole.

CAPITOLO 3: UTILIZZO DEI BIG DATA NEL MARKETING: ALCUNI CASI ESEMPLIFICATIVI

Dopo aver esaminato in generale come le aziende si sono approcciate alla rivoluzione dei Big Data, in questo capitolo verranno analizzati nello specifico casi reali di alcune grandi multinazionali leader nei loro rispettivi campi (Facebook per i social media, Google per i motori di ricerca, Netflix per la distribuzione di contenuti audiovisivi in streaming, Amazon per l'eCommerce, Spotify per la musica on demand) che hanno compiuto un enorme sviluppo in termini di digitalizzazione e incremento degli standard economico-finanziari grazie ai Big Data e alle Big Data analytics.

3.1 Facebook

Facebook è un social media nato nel 2004 negli Stati Uniti, approdato in Europa nel 2008 e classificato dal 2017 come primo servizio attivo di rete sociale per numero di utenti attivi, ben 2.23 miliardi.

Facebook profila i propri utenti analizzando i Big Data da loro generati (i commenti sui post, le foto pubblicate, i "mi piace", i contenuti condivisi, ecc), determinandone il comportamento tramite i seguenti strumenti¹⁵:

- Cookie di tracciamento: se un utente ha effettuato l'accesso a Facebook e naviga contemporaneamente su altri siti Web, Facebook può tenere traccia dei siti che sta visitando tramite i cookies di tracciamento;
- Riconoscimento facciale: Facebook può tenere traccia dei suoi utenti su Internet e su altri profili della piattaforma con dati di immagine forniti tramite la condivisione degli utenti stessi;

¹⁵ <https://www.simplilearn.com/how-facebook-is-using-big-data-article>

- Analisi dei “Mi piace”: tramite questo tipo di analisi si possono prevedere in modo molto accurato informazioni quali l’orientamento sessuale, l'intelligenza, la stabilità emotiva, la religione, l'uso di alcol e droghe, lo stato delle relazioni, l'età, il sesso, la razza e le opinioni politiche di un utente.

La società statunitense interagisce direttamente con i suoi utenti in diversi modi.

Ad esempio, in occasione del suo decimo anniversario, Facebook ha offerto ai propri utenti la possibilità di visualizzare e condividere un video (“Flashback”), consistente in una raccolta delle foto e dei post che hanno ricevuto il maggior numero di commenti e Mi piace, in cui si ripercorre l'andamento della propria attività sul social. Da allora sono stati creati altri video, ad esempio quello per celebrare un "Friendversary" (l'“anniversario di amicizia” di due persone), e un video speciale nel giorno del proprio compleanno.

Un ulteriore esempio di interazione è la funzione “Ho votato” (“I voted”), mediante cui Facebook ha collegato con successo l'attività politica al coinvolgimento degli utenti, creando un adesivo che consente loro di dichiarare "Ho votato" sui loro profili. È probabile che gli utenti che hanno notato il pulsante abbiano votato e abbiano parlato apertamente del comportamento di voto una volta che hanno visto che i loro amici vi stavano partecipando.

A seguito della sentenza della Corte Suprema Americana sul matrimonio tra persone dello stesso sesso come diritto costituzionale, Facebook ha introdotto la funzione “Celebrare l’orgoglio” (“Celebrate Pride”), un modo per mostrare sostegno all'uguaglianza del matrimonio che consiste nell’applicare un filtro color “arcobaleno” alla propria foto profilo.

La funzione di Facebook più utile agli esperti di marketing di un'azienda, che offre per la prima volta una visione completa e fruibile del proprio pubblico, è sicuramente "Topic Data" ("Dati Argomento"): tale tecnologia mostra le risposte degli utenti su specifici marchi, eventi, attività e argomenti, che vengono poi utilizzate per modificare selettivamente il modo in cui commercializzano sulla piattaforma.

3.2 Google

Google è il motore di ricerca più longevo e usato al mondo, e da oltre 20 anni vede una crescita esponenziale dei propri dati di interazione ed economici. Con l'aiuto di diversi strumenti e tecniche di big data, Google è in grado di esplorare milioni di siti Web e di procurare la risposta o le informazioni giuste in pochi millisecondi.

Per comprendere le esigenze e le preferenze dell'utente, in base ai dati che egli continuamente genera (le ricerche, i preferiti, il numero di click su un sito, la posizione geografica, ecc), Google utilizza¹⁶:

- Pagine indicizzate: sono le pagine Web memorizzate nell'indice di ricerca di Google per rispondere alle query di ricerca. L'indicizzazione implica l'assegnazione di parole o frasi chiave alle pagine Web all'interno di un tag di metadati in modo che la pagina possa essere recuperata facilmente;
- Feed di dati in tempo reale: Google possiede un enorme hub centrale per feed di dati in tempo reale su tutto ciò che è misurabile, come bollettini meteorologici, rapporti di viaggio, mercato azionario e azioni, suggerimenti di acquisto, suggerimenti di viaggio, ecc;

¹⁶ <https://tweakyourbiz.com/technology/google/big-data>

- Strumenti di ordinamento: gli algoritmi di Google eseguono calcoli complessi destinati a far corrispondere le domande inserite dall'utente nella barra di ricerca con tutti i dati disponibili, cercando di determinare se l'utente sta cercando notizie, persone, fatti o statistiche e recuperare i dati dal feed appropriato;
- Pagine Knowledge Graph: Google Knowledge Graph è un database che raccoglie tutti i dati e i fatti su persone, luoghi e oggetti, e viene successivamente utilizzato per fornire risposte utili alle domande degli utenti in modo semplice e rapido;
- Ricerca letterale e semantica: lo scopo principale del motore di ricerca letterale è trovare la radice della frase, cercando una corrispondenza per parte della parola o per l'intera frase. La radice della frase viene quindi esaminata ed esplorata per visualizzare risultati di ricerca migliori;
- Cookie di monitoraggio: servono a Google per tenere traccia dei siti che l'utente sta visitando. Consentono pertanto di raccogliere diversi dati relativi agli utenti come le loro preferenze, le loro inclinazioni, ecc;
- Google+: nel momento in cui si accede al proprio account, Google utilizza la cronologia delle ricerche, le tendenze e la posizione per fornire risultati di ricerca accurati. A questo scopo, Google raccoglie tutti i dati relativi alla frequenza dei siti visitati, le frasi di ricerca utilizzate, i tempi, i dati scaricati, ecc;
- Sinonimi: le frasi vengono comprese attraverso un sistema che ne analizza la radice e la relazione in base alla cronologia delle ricerche passate, alle tendenze e alle relazioni reciproche;
- Google Traduttore: il servizio di traduzione di Google analizza milioni di altri testi o discorsi tradotti per determinare l'interpretazione più precisa;

- Google AdWords: Ogni volta che l'utente naviga in diversi siti Web, l'azienda apprende le sue preferenze, simpatie, antipatie, inclinazioni (ecc), in base alle quali poi Google, tramite il servizio 'AdWords', mostra diversi annunci relativi a prodotti o servizi che potrebbero interessargli.

3.3 Netflix

Netflix è una delle più grandi aziende leader nella distribuzione di contenuti online, principalmente serie TV e film, con oltre 180 milioni di utenti attivi nel mondo. Raccogliendo dai loro abbonati i dati di interazione e risposta ad uno spettacolo televisivo (ad esempio, l'ora e la data in cui un utente ha guardato uno spettacolo, il dispositivo utilizzato, se lo spettacolo è stato messo in pausa, se lo spettatore riprende a guardare dopo la pausa, quanto tempo impiega un utente a finire un programma, il numero di ricerche e cosa viene cercato, ecc)¹⁷ e implementando modelli di analisi al fine di scoprire il loro comportamento e i loro modelli di acquisto, consiglia film e programmi TV in base alle loro preferenze utilizzando come indicatore la percentuale di compatibilità.

Questo approccio per la creazione di contenuti basato sui Big Data ha un tale successo che, rispetto alla media dell'industria televisiva, dove solo il 35% degli spettacoli viene rinnovato dopo la prima stagione, Netflix rinnova il 93% delle sue serie originali.

L'algoritmo di raccomandazione della piattaforma alimenta il 75% dell'attività degli spettatori; in altre parole, tre quarti di tutta la visione di Netflix si basa sui propri

¹⁷ <https://seleritysas.com/blog/2019/04/05/how-netflix-used-big-data-and-analytics-to-generate-billions/>

suggerimenti. Inoltre, Netflix ottimizza anche l'artwork dei singoli film per ogni utente, al fine di evidenziare gli aspetti del film più rilevanti per ognuno di loro.

Il sistema di feedback della piattaforma contribuisce alla personalizzazione dei contenuti per l'utente: in precedenza, gli utenti valutavano film e serie TV con 1-5 stelle, ma nel 2017 i team di prodotto, dopo una lunga serie di test, hanno scoperto che era più semplice ed efficace un sistema di valutazione "pollice su/giù".

In alcune occasioni, Netflix sceglie anche di "mettere l'utente al posto di guida": nel 2018, infatti, la piattaforma ha pubblicato un episodio di *Black Mirror* in cui gli spettatori potevano scegliere il finale. Questo tipo di contenuto high-tech "scegli la tua avventura" porta a un livello completamente nuovo di coinvolgimento.

Storicamente, il lancio di una serie implicava prendere decisioni basate su valutazioni storiche e "sentimenti istintivi". Inevitabilmente, centinaia di serie TV sono state cancellate entro poche settimane dal loro lancio a causa del basso numero di spettatori e delle recensioni negative. Senza accesso a dati concreti sul comportamento degli spettatori, i dirigenti spesso facevano ipotesi sbagliate che non si allineavano con ciò che il mercato voleva, e ciò ha rappresentato un costo enorme per gli editori in termini di sviluppo e marketing.

Poiché Netflix ha una visione dei comportamenti di milioni di abbonati in tutto il mondo, ha interrotto il tradizionale processo di "via libera" a nuove serie TV e film semplicemente osservando ciò che le persone stanno già guardando e come lo stanno guardando. *House of Cards*, ad esempio¹⁸, non ha avuto il via libera solo perché

¹⁸ <https://medium.com/swlh/how-netflix-uses-big-data-to-build-mountains-of-money-829364caefa7>

sembrava una buona trama: la decisione si è basata su una serie di variabili, affidandosi quasi interamente ai dati. Utilizzando le metriche combinate raccolte attraverso la loro piattaforma, Netflix ha determinato che una percentuale significativa dei suoi abbonati aveva guardato in streaming il film del regista David Fincher, *The Social Network*, dall'inizio alla fine, e che i film con Kevin Spacey hanno sempre avuto successo con il suo pubblico.

3.4 Amazon

Amazon è l'azienda numero uno al mondo nell'eCommerce: nasce nel 1994 come libreria online, ma deve il suo sviluppo in concomitanza a quello del settore eCommerce nella seconda decade degli anni 2000.

Amazon raccoglie un'ampia quantità di dati su come gli utenti interagiscono con i vari elementi del sito (le ricerche, i preferiti, i prodotti aggiunti al carrello, ecc), monitorando il loro percorso dalla scoperta dei prodotti all'acquisto e alla consegna finale o alla restituzione, per fornire loro la migliore esperienza possibile, nonché perfezionare continuamente le prestazioni del sito. Dal lato dei clienti, utilizza sofisticati algoritmi per far emergere i prodotti e i venditori più pertinenti; dal lato dei venditori, fornisce dati e metriche che possono aiutarli a gestire le loro operazioni, ottimizzando il modo in cui visualizzano le informazioni sul sito o conducono la pubblicità.

Amazon utilizza l'analisi predittiva per un marketing mirato, che possa aumentare la soddisfazione e la fedeltà dei clienti, applicandola a diversi strumenti¹⁹.

¹⁹ <https://www.analyticssteps.com/blogs/how-amazon-uses-big-data>

Amazon offre assistenti virtuali a comando vocale come Echo, Echo Show ed Alexa. I files audio delle registrazioni vocali, caricati poi sui propri server, aiutano Amazon a migliorare sempre più le loro funzionalità, permettendo un migliore riconoscimento vocale da un gruppo di clienti diversi tra loro e quindi una più accurata elaborazione dei messaggi.

Amazon è leader nell'utilizzo di motori di filtraggio collaborativo completo per sistemi di raccomandazione personalizzata. L'azienda analizza i comportamenti di acquisto dei clienti dagli articoli precedentemente acquistati, da quelli nel carrello o nella loro lista dei desideri, dai prodotti da loro recensiti e valutati e da quelli più cercati. Queste informazioni vengono quindi utilizzate per consigliare prodotti aggiuntivi che altri clienti hanno acquistato al momento dell'acquisto degli stessi articoli.

Il modello di spedizione anticipata di Amazon utilizza le informazioni di ogni cliente per prevedere gli articoli che probabilmente acquisteranno, quando potrebbero ottenerli e dove potrebbero richiedere gli articoli: gli oggetti vengono spediti da un centro di circolazione del quartiere o da un magazzino in modo che vengano preparati per il trasporto una volta che il cliente le richiede.

Nel 2013, dopo aver acquistato Goodreads, Amazon ha integrato il servizio di social network dei suoi utenti con la funzione "Kindle Highlighting" di Kindle: ciò ha permesso agli utenti di evidenziare parole e note e anche di condividerle tra di loro, come modo per discutere del libro. In questo modo l'azienda può monitorare regolarmente le parole evidenziate in Kindle al fine di conoscere l'interesse dei lettori, consigliare loro altri eBook e migliorare l'esperienza di lettura.

Inoltre, Amazon utilizza le informazioni estratte dai Big Data generati dai consumatori anche per:

- Ottimizzare la catena di fornitura: per evadere rapidamente gli ordini, Amazon si connette con i produttori e tiene traccia dei loro inventari, analizza i dati disponibili e individua il magazzino più vicino a un cliente per ridurre i costi di spedizione;
- Ottimizzare i prezzi: i prezzi su Amazon sono calcolati in base alla disponibilità dell'utente ad acquistare, ai prezzi della concorrenza, alla disponibilità del prodotto, alle preferenze degli articoli, alla cronologia degli ordini, al margine di profitto previsto e ad altri fattori. I prezzi dei prodotti generalmente cambiano ogni minuto, man mano che i Big Data vengono aggiornati e analizzati. Di conseguenza, Amazon offre in genere sconti sugli articoli più venduti e guadagna maggiori profitti sugli articoli meno popolari;
- Esaminare gli acquisti e le richieste di reso: per evitare frodi al dettaglio, l'azienda raccoglie migliaia di dati storici e in tempo reale su ogni ordine e utilizza algoritmi di apprendimento automatico per trovare transazioni con un'elevata probabilità di essere fraudolente e le richieste di reso dubbie;
- Modificare i negozi fisici: nonostante l'azienda non abbia realmente esposto le informazioni che raccoglie sui clienti di Amazon Go (il supermercato ipertecnologico senza casse né cassieri, che sfrutta dei sensori per identificare i prodotti che i clienti prendono dagli scaffali e che permette di pagare tramite smartphone) e il motivo, è possibile che le utilizzi per migliorare i propri negozi. Ad esempio, se le telecamere hanno indicato che le persone con passeggini hanno

avuto difficoltà a esplorare le passerelle, Amazon potrebbe intervenire e renderle più estese.

3.5 Spotify

Spotify, la più grande piattaforma di streaming musicale al mondo, incorpora intelligenza artificiale, tecniche di apprendimento automatico e Big Data allo scopo di offrire agli utenti un'esperienza di ascolto personalizzata ed esclusiva²⁰.

Spotify utilizza i dati generati da ogni utente (quali canzoni ha ascoltato, quante volte è stata ascoltata una determinata canzone, quali sono gli artisti più ascoltati, ecc) per suggerire contenuti esclusivi per i loro gusti. L'obiettivo è garantire loro un'esperienza soddisfacente in modo che diventino clienti duraturi. Un ruolo fondamentale nella raccolta dati è svolto dalla funzione "Discover", emersa come playlist delle canzoni rilasciate dagli artisti amati dall'utente, ma che si è lentamente sviluppata in una sorta di motore di raccomandazione che suggeriva una raccolta di tracce allineate ai brani contenuti nella playlist. L'algoritmo esamina le playlist di altri utenti per determinare le somiglianze tra le tracce e quindi adotta tali dati per sviluppare una nuova playlist che si allinei con le preferenze di traccia prevalenti dell'utente. Al fine di poter personalizzare queste playlist, la piattaforma ha dovuto prestare molta attenzione sia alle tracce che gli utenti trasmettono in streaming, sia al modo in cui interagiscono generalmente con ogni traccia. Ad esempio, se un brano è stato riprodotto ma è stato saltato in meno dei 30 secondi iniziali, viene percepito come una reazione poco entusiasta e le informazioni sul brano non vengono incorporate durante il calcolo delle playlist. Invece, il fatto che un

²⁰ <https://www.analyticssteps.com/blogs/how-spotify-using-big-data>

brano sia stato aggiunto dall'utente alla propria libreria o playlist e sia stato ascoltato completamente, viene percepito come una reazione positiva, e conferma che il brano ha concordato con il gusto dell'utente.

Il profilo di gusto quotidiano dell'ascoltatore è incorporato anche nelle playlist "Daily Mixes", diverse dai generi musicali verso cui l'utente tende normalmente e generalmente composte sia da brani che l'utente ha salvato o che ha aggiunto alle sue playlists, sia da brani di artisti o album con cui l'utente non ha familiarità. Un altro esempio è quello di "Release Radar", una playlist settimanale che incorpora varie nuove uscite degli artisti che ogni utente segue.

La funzione "Spotify Wrapped" è molto utilizzata a dicembre, in quanto offre agli utenti una carrellata delle loro canzoni preferite ascoltate nell'intero anno. Gli utenti riceveranno un rapporto che li informa se appartengono all'1% dei seguaci più fedeli di una band o dei principali ascoltatori di brani non mainstream. Attraverso questi dati, la piattaforma sviluppa un'esperienza raccontando una storia usando dati musicali, riuscendo a far sentire gli utenti riconosciuti e a suscitare il loro entusiasmo.

Mediante i dati generati dai propri utenti, Spotify è stata anche in grado di aggiornare le proprie campagne pubblicitarie e indirizzare i clienti in maniera più convincente.

L'aggiornamento continuo del proprio sistema permette a Spotify di tenere il passo dei Big Data: ad esempio, all'inizio del 2018, la piattaforma dichiarò che agli utenti gratuiti sarebbe stata concessa la possibilità di esplorare 15 delle sue famose playlist della piattaforma. Ciò ha consentito di raccogliere i dati di oltre cento milioni di utenti aggiuntivi.

Nel tentativo di rendere disponibile l'enorme quantità di dati agli artisti e ai loro manager, la piattaforma ha introdotto l'applicazione Spotify for Artists, in cui viene fornito l'accesso ad informazioni come, ad esempio, quali playlist hanno contribuito a generare nuovi utenti e il numero di flussi che stanno ricevendo complessivamente. Ciò consente inoltre agli artisti di ottenere un maggiore controllo della loro presenza su Spotify, ad esempio per la scelta della "raccolta dell'artista" e anche per attività come l'aggiornamento della propria biografia o la pubblicazione di playlist.

CONCLUSIONI

L'avvento dei Big Data ha rappresentato una svolta epocale, dalla quale non si torna più indietro. Questa nuova enorme massa di dati generati a velocità inimmaginabili ha investito, più o meno fortemente, tutti i settori della società: dalla comunicazione alla sanità, dalla finanza alla logistica, passando per l'energia e l'agricoltura (ecc). La rivoluzione dei Big Data ha comportato una nuova concezione di dato, inteso come risorsa indispensabile per poter aumentare l'efficienza dei processi decisionali e operativi, nonché diminuirne l'incertezza.

Il cambiamento è stato talmente netto che si è anche assistito alla nascita di una nuova disciplina, la c.d. "data science", e di una nuova figura professionale, quale il "data scientist", colui che sa lavorare coi dati e che da essi sa ricavare un vantaggio competitivo, una figura professionale che in Italia (ma in generale nel mondo) è sempre più ricercata. È pur vero che, essendo un cambiamento di grandissima portata, non tutte le aziende riescono a adattarsi rapidamente per tenere il passo dei dati: il segreto del successo sta nella corretta allocazione delle risorse di capitale fisico, umano e organizzativo, unita ad una dose di "ignoranza creativa" e ad una combinazione di capacità dinamica ed adattiva.

Anche il marketing è stato pesantemente rivoluzionato dall'esplosione a livello mondiale dei Big Data: per i manager diventa ora indispensabile affidarsi ad esperti di marketing che sappiano estrarre dati dalle diverse fonti esistenti (social media, Web, sensori, IoT, sondaggi, ecc) ed utilizzare efficacemente tutti gli strumenti di gestione dei Big Data a disposizione dell'azienda (intelligenza artificiale, data mining, machine learning, regressioni, associazione, classificazione, topic modelling, collaborative filtering, ecc).

Per essere realmente competitiva nel mercato, un'azienda deve poter implementare al proprio interno strumenti in grado di effettuare analisi descrittive sulla situazione dei processi aziendali e su altri fenomeni d'interesse, azioni diagnostiche che supportino le ipotesi da testare, analisi predittive per un maggior controllo della gestione e analisi prescrittive per risolvere eventuali problemi riscontrati.

Le aziende che riescono ad adattare la loro struttura e i loro processi a questo cambiamento hanno sicuramente maggiori probabilità di conseguire un vantaggio competitivo sostanzioso e sostenibile: le aziende riportate come esempio nel terzo capitolo hanno tratto un vantaggio incommensurabile (in termini di fatturato) dall'uso dei dati per un marketing mirato, che sappia fornire al cliente un'esperienza di acquisto personalizzata che lo faccia sentire unico.

Entro il 2025²¹, si prevede che oltre 150 trilioni di gigabyte di dati avranno bisogno di essere analizzati. Per il marketing, ciò significa che con il passare del tempo aumenterà la necessità di strumenti avanzati per analizzare e archiviare i dati: molto presto tutti gli sforzi e gli investimenti di marketing andranno concentrati in nuovi software online specificamente orientati alla gestione dei big data, altrimenti gli strumenti odierni risulteranno, con ogni probabilità, decisamente obsoleti.

Ottenere un vantaggio competitivo nel proprio campo diventerà più difficile poiché più esperti di marketing di diverse aziende avranno accesso agli stessi dati sui consumatori: ciascuno dovrà necessariamente distinguersi dagli altri attraverso l'innovazione.

²¹ <https://www.forbes.com/sites/forbesagencycouncil/2020/01/17/the-past-present-and-future-of-big-data-in-marketing/?sh=699197405da9>

BIBLIOGRAFIA

DELMASTRO M., NICITA A. - Big Data – Come stanno cambiando il nostro mondo Il Mulino, Bologna 2019.

GREGORI G.L., PASCUCCI F. - Il digital marketing come fattore competitivo. Verso un approccio integrato «strumenti» e «strategia», 2019.

IANDIORIO E. - Big Data – Cosa sono, come analizzarli e utilizzarli per fare marketing, Hoepli, Milano 2019.

MANDELLI A. – Big Data Marketing - Creare valore nella platform economy con dati, intelligenza artificiale e IoT, EGEA Bocconi, Milano 2017.

MARCELLO M. MARIANI, SATISH NAMBISAN - Innovation Analytics and Digital Innovation Experimentation: The Rise of Research-driven Online Review Platforms, Technological Forecasting and Social Change, Volume 172 (2021).

SAMUEL ISRAEL GOYZUETA RIVERA - "Big Data Marketing: un approccio" (2015) pp. 147-158.

SHAOKUN FAN, RAYMOND Y.K. LAU, J. LEON ZHAO - Demystifying Big Data Analytics for Business Intelligence Through the Lens of Marketing Mix (2015) pp. 28-32.

SHIVAM GUPTA, THÉO JUSTY, SHAMPY KAMBOJ, AJAY KUMAR, EIVIND KRISTOFFERSEN - Big Data and firm marketing performance: Findings from knowledge-based view, Technological Forecasting and Social Change, Volume 171 (2021).

SUNIL EREVELLES, NOBUYUKI FUKAWA, LINDA SWAYNE - Big Data consumer analytics and the transformation of marketing, Journal of Business Research, Volume 69, Issue 2, Pages 897-904 (2016).

WEDEL M, KANNAN P.K. - Marketing Analytics for Data-Rich Environments (2016) pp. 97-121.

SITOGRAFIA

<https://link.springer.com/article/10.1007/s00500-017-2536-4>

<https://www.themarketingfreaks.com/2019/11/big-data-cosa-sono-la-storia-le-caratteristiche-le-analisi-ed-esempi/>

<https://www.digital4.biz/marketing/social-media-analytics-cos-e/>

<https://www.digital-coach.it/blog/case-histories/social-media-analytics/>

<https://www.digital-coach.it/web-analytics/>

<https://www.tsw.it/journal/archivio/i-fondamentali-del-web-marketing-la-web-analytics/>

<https://www.studiosamo.it/glossario/analytics/>

<https://www.criteo.com/it/what-is-retargeting/>

<https://www.zerounoweb.it/analytics/retail-analytics-larma-segreta-dellit-store-manager-per-conquistare-business-e-clienti/>

<https://www.buxtonco.com/blog/how-mobile-gps-data-can-take-you-where-you-want-go>

<https://www.digital4.biz/whitepapers/big-data-analytics-trend-e-tecnologie-su-cui-puntare-per-migliorare-il-business/>

<https://seleritysas.com/blog/2019/04/05/how-netflix-used-big-data-and-analytics-to-generate-billions/>

<https://medium.com/swlh/how-netflix-uses-big-data-to-build-mountains-of-money-829364caefa7>

<https://www.analyticssteps.com/blogs/how-amazon-uses-big-data>

<https://rehack.com/data/5-ways-how-amazon-uses-big-data/>

<https://www.analyticssteps.com/blogs/how-spotify-using-big-data>

<https://tweakyourbiz.com/technology/google/big-data>

<https://www.simplilearn.com/how-facebook-is-using-big-data-article>

<https://www.aliven.it/2019/06/07/i-big-data-nella-gestione-della-catena-di-distribuzione/>

<https://www.mecalux.it/blog/big-data-logistica>

<https://www.osservatori.net/it/prodotti/formato/report/la-figura-del-data-scientist-e-l-impatto-a-livello-organizzativo>

<https://www.osservatori.net/it/prodotti/formato/report/trasformazione-data-driven-in-atto-competenze-cultura-analytics-report>

<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/forbesagencycouncil/2020/01/17/the-past-present-and-future-of-big-data-in-marketing/?sh=699197405da9>