



Dipartimento di Impresa e Management.

Cattedra di Economia e Gestione delle Imprese.

BIG DATA:
UNA RIVOLUZIONE PER LE IMPRESE NEL SEGNO DELL'
INNOVAZIONE

RELATORE

Prof.ssa Maria Isabella Leone

CANDIDATO

Vincenzo Cappelletti

Matricola: 213661

ANNO ACCADEMICO 2018-2019

*A mio padre,
tutta un'adolescenza.*

Sommario

L'UNIVERSO DEI DATI.....	7
1.1. BREVE INTRODUZIONE.....	7
1.2. CENNI STORICI.....	8
1.3. CARATTERISTICHE DEI BIG DATA.....	10
1.4. LA CATENA DEL VALORE DEI BIG DATA.....	13
1.5. ETICA E LEGISLAZIONE.....	17
BIG DATA PER IL MARKETING.....	21
2.1 BIG DATA ADVERTISING	21
2.2 INTRODUZIONE AL PROGRAMMATIC ADVERTISING:.....	23
2.3 COME FUNZIONA IL PROGRAMMATIC ADVERTISING:	24
2.4 FACEBOOK PIONIERE DEL PROGRAMMATIC ADVERTISING	26
2.5 CONSUMER-CENTRIC PROGRAMMATIC ADVERTISING	27
2.5.1 L'EVOLUZIONE DEL PROGRAMMATIC ADVERTISING	27
2.5.2 CONSUMER CENTRIC MARKETING: UNA VISIONE DEL SINGOLO CONSUMATORE.....	30
2.6 RISCONTRI, RAZIONALITÀ E VALORE	32
3 DIGITAL PLATFORM E BIG DATA	34
3.1 I BIG DATA GUIDANO LA RIVOLUZIONE DEI BUSINESS	34
3.2 IN CHE MODO I BIG DATA ALIMENTANO LA SHARING ECONOMY	36
3.3 IL CASO AIRBNB	39
3.4 BARRIERE ALL' INGRESSO NELLA PLATFORM ECONOMY	40
3.5 I BIG DATA COME FONTE DI "PRICE DISCRIMINATION"	43
3.6 L'EFFETTO DEI BIG DATA SULL' INNOVAZIONE.....	45
3.7 ABUSI E POSIZIONI DOMINANTI	47
POLITICAL MICROTARGETING: IL CASO CAMBRIDGE ANALYTICA	49
4.1 LA PROFILAZIONE DELL'UTENTE	49
4.2 POLITICAL MICROTARGETING	51
4.3 MICROTARGETING POLITICO ATTUATO DA CAMBRIDGE ANALYTICA	53
4.4 IL CASO CAMBRIDGE ANALYTICA	54
4.5 SVILUPPI DELLA VICENDA E REGOLAMENTAZIONE.....	58

4.6	SANZIONI APPLICABILI PER LA VIOLAZIONE DEL REGOLAMENTO	62
4.7	PANORAMA ATTUALE	63
CONCLUSIONE		64
BIBLIOGRAFIA		66
SITOGRAFIA		68

Introduzione

“*data is the new oil*” i dati sono il nuovo petrolio, diceva nel 2006 il matematico Clive Humby, intorno a questa frase poi nel 2017 la rivista *The Economist* ha pubblicato un’articolo intitolato “*The world's most valuable resource is no longer oil, but data*” con il chiaro intento di accendere i riflettori sulla rivoluzione apportata dai dati, un fenomeno di difficile lettura anche a causa della sua continua volubilità. Il parallelismo con i giganti del petrolio occorso da Huby era calzante, ma forse riduttivo in uno scenario che vede i nuovi “estrattori” di dati padroni di una risorsa assai strategica.

Le *big Five* del *tech* Alphabet, Amazon, Apple, Facebook e Microsoft hanno raggiunto fatturati paragonabili a quelli di intere nazioni. Tutte assieme le “5 stelle” citate precedentemente valgono oggi sul mercato quanto l’intero Pil della maggiore potenza europea, ossia la Germania. Nel primo trimestre 2018 i colossi del *tech* hanno collezionato entrate da *record*, la sola Apple ha raggiunto per la prima volta il traguardo dei venti miliardi di utili trimestrali, Amazon ha registrato profitti per quasi due miliardi di dollari aumentando le entrate del trentotto per cento ed oggi possiede anche interi studi cinematografici ad hollywood dove girare serie televisive e film, supermercati e si prepara a lanciare la sua sfida al mondo della sanità collaborando con Berkshire Hathaway e JP Morgan.

Il lavoro inizia dunque proprio con la disamina dell’universo big data nel primo capitolo. Partendo dai cenni storici, l’estratto tratta poi quelle che sono le principali caratteristiche che identificano i big data, successivamente l’attenzione si sposta sulla catena del valore particolarmente apprezzata dal mondo delle industrie per inquadrare tale fenomeno all’interno delle stesse, il primo capitolo volge quindi al termine con una riflessione sugli aspetti etico-legali.

Il secondo capitolo sposta lo sguardo sul rapporto tra l’impresa ed il consumatore, ed in particolar modo come questo sia cambiato con l’avvento dei big data. Un primo paragrafo ruota attorno all’*advertising* ed è seguito da un secondo e un terzo in cui vengono approfonditi prima il cosiddetto *programmatic advertising* e successivamente le modalità di funzionamento dello stesso, nel quarto paragrafo si illustrano le implicazioni pratiche

di tale fenomeno all' interno dell'azienda Facebook. Nel quinto paragrafo viene studiata l'evoluzione del *programmatic advertising*, che prende il nome di *consumer-centric programmatic advertising*. Il capitolo nel suo volgersi alla conclusione sottolinea i riscontri sul valore delle nuove pratiche di *advertising* rilevato da parte delle aziende.

Nel penultimo capitolo si vuole spiegare come i big data abbiano contribuito alla rivoluzione delle piattaforme digitali. Il primo paragrafo ha proprio l'intento di mostrare come la ricchezza di informazioni di cui sono in possesso le imprese abbia permesso il sorgere delle *digital platform* con annessi nuovi business. Nel secondo paragrafo si analizza una conseguenza a cui si è giunti grazie alle piattaforme digitali, ossia la *sharing economy* con annesso esempio (caso Airbnb) nel terzo paragrafo. Le conseguenze per le economie con l'avvento dei big data sono affrontate nei successivi quattro paragrafi; dapprima sono analizzate le nuove di barriere all' ingresso, poi il fenomeno della *price discrimination*, successivamente l'effetto dei dati sull' innovazione e nel settimo ed ultimo paragrafo gli abusi e le posizioni dominanti.

L'ultimo capitolo infine prende spunto dal caso Cambridge analytica e Facebook. La trattazione del capitolo comincia con la spiegazione della profilazione dell'utente grazie ai big data; si spiega nel secondo paragrafo la pratica del *microtargeting* politico e nel successivo paragrafo come la società statunitense Cambridge analytica lo abbia utilizzato nella sua attività facendone prima la propria fortuna ma dando poi il via a una lunghissima controversia legale sulla privacy. Infine gli ultimi tre paragrafi descrivono come sia cambiato il panorama legislativo dopo i recenti scandali legati all' uso improprio dei dati degli utenti.

CAPITOLO PRIMO

L'UNIVERSO DEI DATI

1.1. BREVE INTRODUZIONE

In un'economia sempre più globale e dinamica, le organizzazioni hanno cominciato a fare sempre più affidamento sulle intuizioni dei propri clienti, al fine di scoprire nuove opportunità di crescita. Il processo di scoperta e determinazione di queste opportunità, è generato da insiemi di dati complessi che devono essere gestiti, analizzati e manipolati da professionisti qualificati. Il fenomeno in questione, che consente questa raccolta massiccia di informazioni è noto come *big data*. La maggior parte di questi dati viene raccolta in maniera automatica grazie a sistemi di *digital analytics* (*web, social, app, in-store*), grazie soprattutto alla rete web anche mobile in grado di tracciare i comportamenti degli utenti. Ma sono le macchine stesse a rappresentare la fonte in più rapida ascesa per la raccolta di dati, mediante i sensori all' interno dei nostri *device*. Perché i dati possano essere considerati realmente *big* devono essere generati in tempo reale, ossia godere dell'attributo *real time generated*, attributo che ha come caratteristica principale quella di minimizzare al massimo il tempo che separa la ricerca del dato dal momento in cui viene utilizzato per la decisione. Per rendere meglio l'idea, basta pensare al flusso di dati che utilizzano le aziende gestori delle carte di credito per monitorare ed identificare tempestivamente eventuali frodi. Quando viene rilevata una piccola spesa e successivamente viene utilizzata la carta per una grande spesa si costituisce uno *streaming event* che avvisa la necessita di un provvedimento anche automatico (avviso al cliente e/o sospensione momentanea della carta).

Definire la portata in termini quantitativi di questo fenomeno è una sfida in continuo cambiamento, l'università di Villanova in un recente studio afferma che per avere un'idea della portata di tale fenomeno bisogna prendere in considerazione le seguenti informazioni degne di nota:

Secondo IBM, gli utenti creano 2,5 quintilioni di byte di dati ogni giorno. In termini pratici, ciò significa che il 90% dei dati nel mondo di oggi è stato creato solo negli ultimi due anni *Walmart* controlla più di 1 milione di transazioni dei clienti ogni ora, che

vengono poi trasferite in un database che funziona con oltre 2,5 *petabyte* di informazioni. Secondo FICO, il sistema di contrasto dalle frodi con carta di credito consente di proteggere oltre due miliardi di account in tutto il mondo *Facebook* detiene più di 45 miliardi di foto nel suo database di utenti, un numero che cresce ogni giorno. Questi dati sono così big che già quelli raccolti fino al 2014 se fossero stati tutti stampati su fogli di carta A4 avrebbero creato una pila così alta da percorrere la distanza tra la Terra e Giove quattordici volte (A.mandelli, 2017).

1.2. CENNI STORICI

“C’è chi dice che i big data sono nati quando il costo di raccogliere e gestire un’informazione è diventato inferiore al costo di cancellarla”. (A.mandelli, 2017)
 La storia dei big data è strettamente collegata alla possibilità di archiviare e gestire in maniera sempre più efficiente quantità di dati sempre maggiori, nello specifico quanto più tale processo viene migliorato, tanto più sarà possibile assistere allo sviluppo di nuove *database technologies*

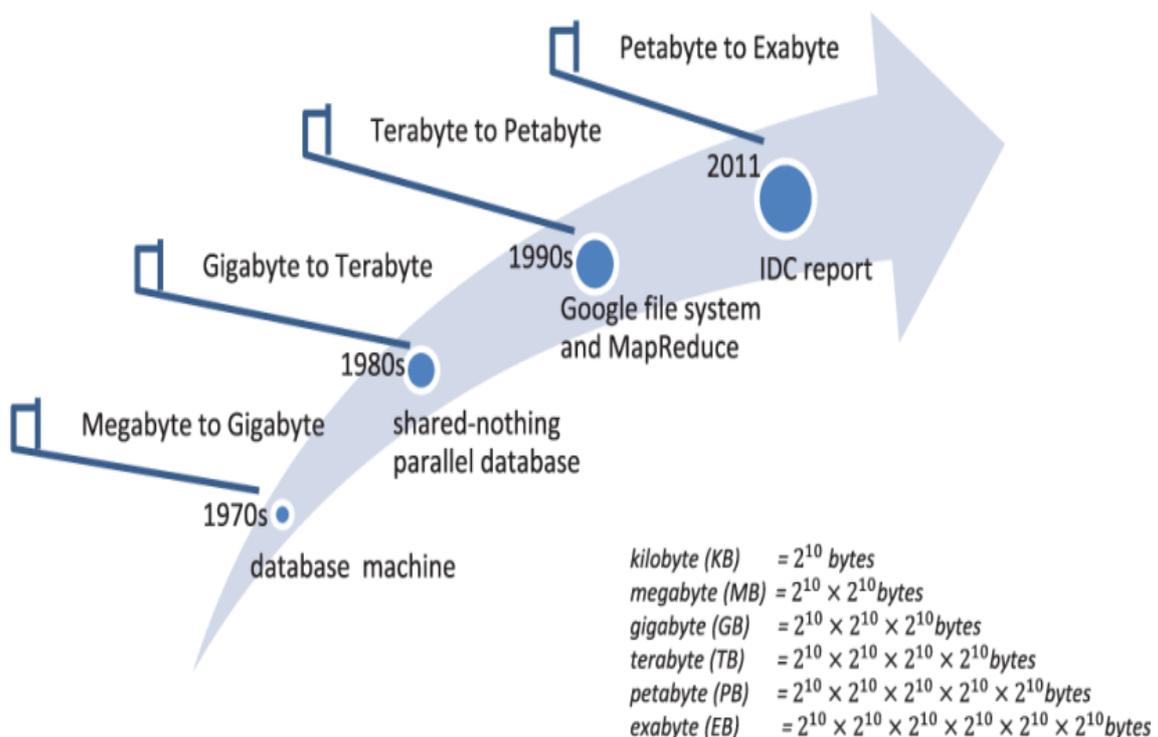


Fig. 11

Le fasi salienti di questo percorso affondano le radici negli anni '60 con la nascita del primo *Main Frame Computer* di enormi dimensioni in grado di digitalizzare e archiviare i dati, non era ancora possibile condividere questi dati e per accedervi bisognava recarsi fisicamente sul posto. Tra gli anni '70 e gli anni '80 la storia dei big data è stata oggetto di una grande rivoluzione che vede come protagonista il passaggio da *megabyte* a *gigabyte*. A quel tempo la spinta era data dalla necessità di gestire ed immagazzinare grandi quantità di dati soprattutto ai fini di analisi aziendali. La ricerca si spinse sino ad elaborare un'integrazione tra *hardware* e *software* in grado di poter lavorare in maniera più efficiente ed allo stesso tempo ad un costo minore. Tuttavia divenne chiaro che la creazione di tale macchina fosse fortemente penalizzante in termini sia di costo che di progresso tecnologico in quanto risultò molto più conveniente sviluppare dei *software* che si adattassero ai generici computer che crescevano a livello capacitivo in maniera esponenziale. Pertanto, i sistemi di database discendenti da questo processo sono sistemi software che impongono pochi vincoli all'hardware e possono essere eseguiti su computer generici. Alla fine degli anni '80 la tecnologia digitale divenne sempre più di uso comune tanto da portare i volumi di dati ad espandersi fino a diversi *gigabyte* o addirittura a un *terabyte*, ben oltre le capacità di archiviazione e / o di elaborazione di un singolo grande computer. Ciò ha comportato la *shared-nothing architecture* ossia la costruzione di reti articolate di computer singoli ed indipendenti ma allo stesso tempo collegati tra loro, aventi lo scopo di estendere le capacità di archiviazione e migliorare le prestazioni distribuendosi i dati e le attività correlate.

Negli anni '90 il rapido sviluppo del web portò l'intero mondo all'era di internet, accompagnato da un'elevata quantità di pagine web strutturate e semi strutturate contenenti *terabytes* o *petabytes* di dati. Per affrontare la sfida della gestione e dell'analisi dei dati su

¹Han Hu, Yonggang Wen, Tat-Seng Chua, Xuelong Li, 2014 Toward Scalable Systems for Big Data Analytics: A Technology Tutorial breve storia dei big data con importanti pietre miliari. Può essere divisa in quattro fasi in base alla crescita dell'ordine di grandezza dei dati, Megabyte, Gigabyte, Terabyte, Petabyte e Exabyte.

scala web Google creò *Google file system* in grado di ordinare e processare automaticamente i dati.

In epoca recente la grande mole di dati archiviati e analizzati dalle grandi aziende raggiungerà senza dubbio il PB fino alla grandezza di *exabyte*. Nel giugno 2011, EMC ha pubblicato un rapporto intitolato *extracting value from chaos*²³ dove vengono discussi il concetto di big data e il suo potenziale. Questo rapporto ha acceso l'entusiasmo per i big data nell'industria e nel mondo accademico. Negli anni che seguirono, quasi tutte le aziende dominanti del settore, tra cui *EMC, Oracle, Microsoft, Google, Amazon e Facebook*, iniziarono a sviluppare progetti di *big data*.⁴²

1.3. CARATTERISTICHE DEI BIG DATA

L'espressione Big Data è stata utilizzata per la prima volta negli anni 90 dallo studioso *John Mashey*, il quale la adottò per fare riferimento alla gestione ed all'analisi di una innumerevole quantità di *dataset*.

Tuttavia non essendo riscontrabile una definizione universalmente riconosciuta di *big data* è possibile effettuare uno studio concernente questi ultimi attraverso la descrizione delle caratteristiche più pregnanti. *Doug Laney* analizzò, attraverso il modello di crescita tridimensionale (modello delle "3V")⁴, le prime tre caratteristiche:

- **Volume:** indica la dimensione effettiva di un *dataset* ovvero un qualsiasi "database" che sia tanto grande da non poter essere agevolmente gestito da un normale computer. Inizialmente venivano considerati Big Data i database con dimensioni maggiori di un *terabyte* (1.000 gigabyte). Tuttavia, oggi, l'ampiezza ed il volume dei Big Data

² Per approfondimenti si rimanda al report di John Gantz e David Reinsel, *Extracting Value from Chaos* <https://www.emc.com/collateral/analyst-reports/idc-extracting-value-from-chaos-ar.pdf>

³ Per approfondimenti si rimanda all' articolo di Han Hu, Yonggang Wen, Tat-Seng Chua, Xuelong Li, 2014 *Toward Scalable Systems for Big Data Analytics: A Technology Tutorial* <https://ieeexplore.ieee.org/document/6842585/authors#authors>

⁴ Per approfondimenti si rimanda all' articolo di Doug Laney, 2001, *3D Data Management: Controlling Data Volume, Velocity, and Variety*. Current business conditions and mediums are pushing traditional data management principles to their limits, giving rise to novel, more formalized approaches,

rappresentano solo all'apparenza un problema: grazie ai sistemi di *cloud* e di virtualizzazione è possibile gestire e raccogliere una grande quantità di dati ed allo stesso tempo usufruire di processi di accesso ed immagazzinamento agevolati

- **Varietà:** I Big data contengono differenti tipologie di dati strutturati e non. I dati strutturati sono dati ben definiti ed organizzati che possono essere rappresentati da numeri o categorie come ad esempio età o il sesso; al contrario i dati non strutturati sono imprecisi, indefiniti e difficili da categorizzare come ad esempio email, blog e conversazioni telefoniche.
- **Velocità:** i dati possono cambiare a diverse velocità a seconda delle caratteristiche che li contraddistinguono. Alcuni dati sono statici come ad esempio il luogo di nascita, altri possono variare occasionalmente come ad esempio l'indirizzo di casa ed altri ancora possono variare istantaneamente come ad esempio il battito cardiaco.

inoltre negli ultimi anni si tende a riconoscere delle ulteriori caratteristiche (Bernard Marr, 2015) quali:

- **Veracità:** i dati spesso possono essere confusi o poco affidabili, con molte forme di big data, la qualità e l'accuratezza sono meno controllabili. I grandi volumi spesso compensano la mancanza di qualità o accuratezza dei dati.
- **Valore:** caratteristica che ci consente di trasformare i dati in nostro possesso in valore. Per questa caratteristica è importante che le aziende sostengano ogni tentativo di raccogliere e sfruttare i big data. Questi ultimi, infatti, consentono alle aziende di poter comprendere e soddisfare al meglio le esigenze dei propri clienti, oltre che agevolare ed ottimizzare i processi interni alle stesse.

Dagli anni 90 in poi numerosi studiosi hanno incentrato la propria analisi sulla ricerca di quelle caratteristiche che avrebbero potuto essere comuni a tutte le tipologie di Big Data. Nel 2014 *Kitchin* condusse uno studio sulla base del quale fu in grado di differenziare tra di loro diverse categorie di dati a seconda del volume e della complessità degli stessi. Le evidenze della sua ricerca hanno portato ad una prima suddivisione tra piccoli dati e dati piuttosto inflessibili. I primi sono frutto di tecniche appositamente elaborate dalla scienza per limitarne la portata, la temporalità e le dimensioni; i secondi, al contrario, sono dati rigidi e voluminosi che comportano talune complessità sia in relazione alla modalità con

le quali si generano sia in relazione alle modalità di gestione degli stessi. Pur trattandosi di dati mediamente voluminosi, *Kitchin* puntualizza che i *Big Data* si differenzia tanto dagli uni quanto dagli altri: si tratta di dati che vengono generati in modo continuativo e repentino, essendo generalmente flessibili e scalabili. Nel 2015, invece, altri studiosi, hanno innescato una diversa tendenza basata sull'analisi del grado di difficoltà e complessità che contraddistingue i diversi processi di elaborazione ed archiviazione di dati. I risultati di tali studi hanno però portato a concludere che non sia possibile, al giorno d'oggi, rintracciare un numero esatto di caratteristiche comuni volte all'identificazione di una categoria unitaria definibile come *Big Data*. Ciò che è concesso affermare è che si tratti però di una categoria amorfa al cui interno è possibile catalogare tutte quelle serie di dati che abbiamo tra di loro delle caratteristiche quantomeno simili. Infatti, se analizzassimo una ad una le diverse tipologie di *Big Data*, sarebbe istantaneamente percettibile il fatto che una piccolissima percentuale di questi ultimi sia in grado di soddisfare tutte le caratteristiche convenzionali sopra riportate.

Ciononostante le applicazioni dei big data sono quasi infinite ed è grazie a queste che negli ultimi anni si è assistito ad una vera e propria rivoluzione nel mondo digitale. La società stessa è in continua evoluzione a causa delle enormi quantità di dati che ci è possibile raccogliere ed analizzare. L'utilizzo più innovativo di questi ultimi lo si può riscontrare oltre che nel settore commerciale anche in ambiti relativi all'assistenza sanitaria ed alla sicurezza permettendo da un lato alle aziende farmaceutiche di poter sviluppare la propria ricerca sulla base di dati reali garantendo così l'efficienza dei trattamenti sanitari e dall'altro consentendo al governo di poter prevenire eventuali attacchi terroristici o crimini internazionali.

Tuttavia, ciò che maggiormente interessa soprattutto ai fini aziendali è proprio la quinta V, ossia quella del valore. Le imprese infatti devono sempre chiedersi in che modo i big data potranno essere utili al fine di creare valore ed allo stesso tempo dovranno eludere il rischio di “annegare” nei propri dati mentre sono all'assetata ricerca di benefici.

1.4. LA CATENA DEL VALORE DEI BIG DATA

Il sistema dei big data è un sistema complesso, che può essere illustrato attraverso la “catena del valore”, generalmente apprezzata nel mondo delle industrie ed utilizzata al fine di analizzare e scomporre il sistema big data in quattro fasi: generazione, acquisizione, archiviazione ed analisi.

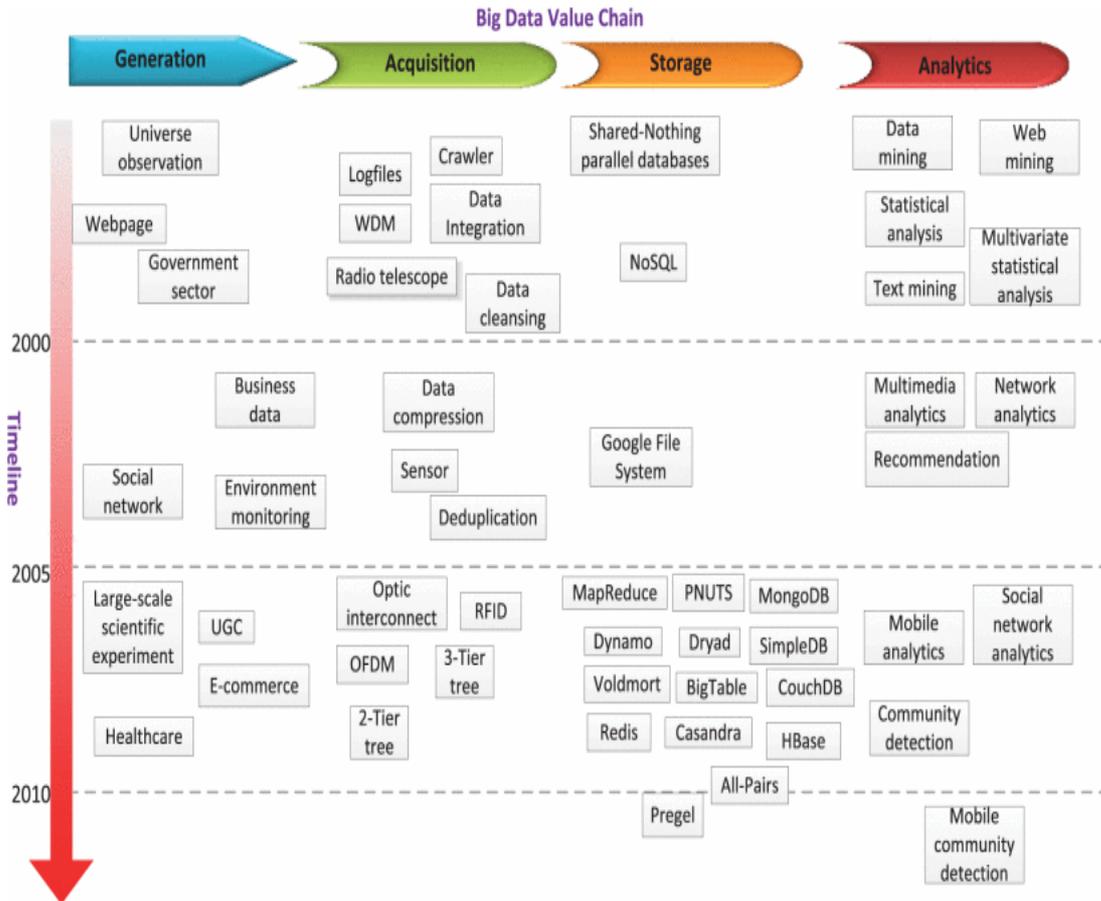


Fig. 2 1

¹ Han Hu, Yonggang Wen, Tat-Seng Chua, Xuelong Li, 2014 Toward Scalable Systems for Big Data Analytics: A Technology Tutorial Mappa della tecnologia dei Big Data. Ruota su due assi, vale a dire la catena del valore dei dati e la timeline. La catena del valore dei dati divide il ciclo di vita dei dati in quattro fasi, tra cui la generazione dei dati, l'acquisizione dei dati, l'archiviazione dei dati e l'analisi dei dati. In ogni fase, evidenziamo tecnologie esemplari negli ultimi 10 anni.

Occorre premettere che, sebbene sia di difficile individuazione, il valore economico è una delle principali caratteristiche dei Big Data soprattutto se l'analisi ha ad oggetto enormi quantità di *dataset* generati sulla base di una varietà di fonti.

In effetti si tratta di veri e propri beni immateriali con caratteristiche distintive come l'infungibilità e la non rivalità: i diversi flussi informativi non sono interscambiabili e possono essere utilizzati da una molteplicità di utenti contemporaneamente. La “catena del valore” è solitamente utilizzata nella realtà aziendale per individuare simultaneamente le attività suscettibili di generare valore e gli eventuali ostacoli legati a tali processi, permettendo alle stesse di eludere le difficoltà legate alla precisa quantificazione del valore di ciascun flusso informativo. Le quattro fasi distintive sono:

- Generazione di dati: in questa fase è illustrata la modalità attraverso la quale sono generati i dati. Tutto ebbe origine negli anni 90 a causa di grandi volumi di dati generati dalle transazioni bancarie e dagli archivi governativi. Successivamente, i motori di ricerca web e le attività di e-commerce hanno contribuito ad accrescere il volume di dati preesistenti attraverso la creazione di una grande quantità di dati semi-strutturati e non strutturati, tra i quali le pagine web. Nel 2000 con la diffusione dei social networks si è arrivati al culmine soprattutto grazie alla possibilità di utilizzarli attraverso dispositivi mobili. Possiamo affermare che si è rapidamente passati da una fase di registrazione passiva di dati ad una di generazione continua ed automatica degli stessi. Ad oggi le fonti principali che generano big data sono *human generated*, *machine generated* e *business generated*. Le prime derivano principalmente dai *social network* (Facebook, Twitter ...), *blog*, *multimedial sharing* (*Instagram*) e dunque, da un processo di espressione dell'essere umano. Le seconde, invece, sono generate da fonti come sensori GPS, dall' *internet of things*, da sistemi di *high frequency trading* ossia da macchine che raccolgono automaticamente dati. Per concludere vi sono le *business generated* che traggono le proprie origini da processi interni alle aziende ad esempio da dati storici, pagamenti ordini, dati di produzione e simili, utilizzati principalmente dalle imprese. I dati *business generated* sono destinati a raddoppiare ogni 1,2 anni.

- **Acquisizione di dati:** in questa fase l'obiettivo è quello di aggregare le informazioni mediante tre *step*: raccolta dei dati, trasmissione e preelaborazione degli stessi. La raccolta dei dati si riferisce al processo di recupero dei dati grezzi dal mondo reale. Tuttavia esistono molti metodi per raccogliere i dati, i principali sono: sensori (acustici, automobilistici, del meteo, della pressione, di prossimità), registri di file come ad esempio quelli dei *server web* che registrano i *click*, gli accessi ecc. ed infine i *web crawler*, ovvero dei software che analizzano i contenuti di un database in maniera automatica per identificare tutti i collegamenti presenti. Per quel che concerne la trasmissione, una volta effettuata la raccolta di dati bisogna trasferirli in una infrastruttura di archiviazione, nota come centro dati. L'ultima fase consiste nella preelaborazione dei dati, poiché, a causa delle diverse fonti da cui otteniamo i *dataset*, questi possono avere diversi livelli di qualità, e pertanto sono state progettate tecniche di preelaborazione che migliorano la qualità dei dati e che ne riducono la ridondanza.
- **Archiviazione di dati:** consiste nel raccogliere ed organizzare le informazioni in un formato conveniente per l'analisi e l'estrazione del valore. L'infrastruttura deve avere due caratteristiche, la prima è quella di poter ospitare le informazioni in modo durevole ed affidabile, la seconda consistente nell'assunto per cui l'archivio deve avere un accesso scalabile per poter consentire l'analisi di una grande quantità di dati.
- **Analisi dei dati:** l'ultima e più importante fase ha l'obiettivo di estrarre valore utili all'azienda e suggerire conclusioni e/o supportare il processo decisionale. Tale intento può essere perseguito attraverso molteplici metodi. Quelli più comuni ed utili a tale scopo sono la *data visualization* che attraverso mezzi grafici fornisce informazioni e l'analisi statistica che attraverso la probabilità cerca di abbattere la casualità e l'incertezza solitamente mediante descrizione ed inferenza.

A tal punto è doveroso specificare come le aziende possono creare valore in ciascuna delle fasi sopra descritte.

In primo luogo è necessario che queste ultime adoperino dei sistemi idonei a ridurre i costi legati all'acquisizione di dati, nonostante le difficoltà legate ad eventuali vincoli sia fisici che legali. Vincoli che nel corso degli ultimi anni hanno subito un'importante

riduzione grazie anche all'utilizzo di canali che via internet consentono la condivisione in tempo reale di dati personali utilizzabili dalle aziende per comprendere quali siano i bisogni che i consumatori intendono soddisfare come l'individuazione della tipologia di pubblicità che meglio sia in grado di sponsorizzare il prodotto di cui questi ultimi necessitano (ad esempio le aziende potrebbero impegnarsi ad offrire un determinato servizio chiedendo come contropartita la condivisione di taluni dati utili per le analisi di mercato). Inoltre per ciò che concerne la problematica dei "consensi" ai fini della trattazione dei dati personali è essenziale che le aziende siano capaci di generare un senso di fiducia negli utenti impegnandosi nella protezione dei flussi informativi condivisi affinché questi ultimi non entrino in possesso di terze parti. Ad un'azienda che sia in grado di soddisfare tali requisiti sarà attribuibile un valore aggiunto in quanto non necessiterà di acquisire le informazioni da fornitori di hardware e/o software che perlopiù comportano lo svantaggio di essere filtrate da terze parti, ma potranno sfruttare le proprie risorse senza dover incorrere in ingenti costi di transazione.

In secondo luogo, in relazione alla fase dell'acquisizione e dell'archiviazione, è necessario che i sistemi elaborati dalle aziende siano capaci di connettere e trasmettere dati anche in tempo reale. In tale ambito i driver di valore sono rappresentati soprattutto dall'affidabilità, dalla sicurezza e dalle prestazioni che la rete è in grado di offrire. Quanti più dati vengono raccolti tanto più saranno elevati i costi ed il tempo che le aziende impiegano per l'archiviazione di questi ultimi ma allo stesso tempo il costo di tale attività sarà in grado di riflettere il valore suscettibile di essere creato nelle fasi successive della catena stessa. Di conseguenza la fase generalmente indicata come quella di "raccolta" di dati acquisisce importanza centrale proprio perché si sostanzia in un evento preparatorio per raccogliere dati e garantire qualità e accuratezza nella fase analitica. Per fortuna, al giorno d'oggi, le aziende, di qualsiasi dimensione esse siano, hanno sviluppato delle tecnologie capaci di archiviare "virtualmente" innumerevoli quantità di dati eludendo in tal modo quelle che in passato potevano rappresentare delle barriere economiche.

Infine la creazione di valore arriva al culmine nella fase dell'analisi. Si tratta di una fase in cui le aziende dovranno impegnarsi nella creazione di strumenti analitici affidabili ed innovativi servendosi, ove possibile, di tecnologie avanzate (come ad esempio

l'intelligenza artificiale). Tali strumenti dovranno pertanto consentire alle aziende non solo di poter analizzare grandi quantità di dati nel minor tempo possibile ma anche di ridurre i costi legati a tali operazioni

1.5. ETICA E LEGISLAZIONE

Agli arbori della *internet economy* il problema relativo alle protezioni dell'individuo concerneva in larga parte la capacità di difenderlo dalle frodi e dall'ingerenza dello spamming. “Da quando le *digital analytics* di tracciamento online e nello spazio geo localizzato sono divenute più sofisticate e i social media si sono evoluti in piattaforme di espressione e condivisione di pezzi di vita personale, il tema della *privacy* è diventato centrale e sempre più gestito come variabile di creazione (o di distruzione) di valore nella relazione con il consumatore” (Mandelli A.,2017).

Il sentimento di forte insicurezza e preoccupazione che creano i big data all' interno della società odierna è ragguardevole, in uno studio svolto negli USA viene mostrato come più di tre quarti dei consumatori ritengono che gli *online advertiser* abbiano eccessive informazioni su di loro e circa metà di questi che i marketer non rispettino le regolamentazioni sulla *privacy* (Michel Wedel & P.K. Kannan, 2016). Inoltre, basti pensare che solo nel 2018 sono stati violati circa un miliardo di account da parte di *criminal hacker*; tutto ciò contribuisce a creare intorno al mondo dei *big data* una giustificata apprensione.

Sebbene si pensi al pericolo dei dati semplicemente soffermandosi sul problema relativo alla *privacy*, tuttavia in realtà è possibile distinguere tre principali motivi di preoccupazione legati all' utilizzo eticamente corretto dei dati personali:

- *Privacy* e proprietà
- Sicurezza ed anonimato
- Processi decisionali

PRIVACY E PROPRIETÀ: una prima visione utilitaristica ritiene che le informazioni ed i dati concernenti un individuo possano essere raccolti e gestiti da chiunque vi abbia accesso. Tale visione comporta aspetti negativi e positivi allo stesso tempo. Da un lato

detenere quante più informazioni possibili per un'impresa vorrebbe dire essere in grado di fabbricare un prodotto su misura del proprio cliente riducendo allo stesso tempo i costi relativi alla gestione del business. Tuttavia l'aspetto negativo riguarda il rischio che tali dati possano essere utilizzati impropriamente come ad esempio nella pratica della discriminazione dei prezzi.

Una seconda visione (*right-based*) che considera i dati come un'estensione naturale della persona afferma che le informazioni personali dovrebbero godere dello stesso rispetto riservato alla persona fisica. Di conseguenza tali dati divengono proprietà personale fornendo il diritto a ciascun individuo di decidere con chi condividere le informazioni personali ed a quali scopi. Tuttavia, potrebbe esserci un disallineamento nel grado di condivisione delle informazioni tra diversi soggetti, comportando una ricaduta negativa sul *direct marketing*. Le imprese, in questo caso, non essendo in possesso di informazioni sufficienti non saranno in grado di fornire ai propri clienti un prodotto specifico. Tali due visioni contrastanti hanno portato a diversi approcci legislativi nei diversi paesi. In alcuni stati con particolare riguardo all' unione europea si è preferito optare per l'approccio *right-based*. La legislazione europea in ambito di *data protection* si basa sui seguenti principi:

- Consenso: è necessario il permesso dell'individuo affinché si possano ottenere ed utilizzare dati personali
- Proporzionalità: i dati di un individuo devono essere utilizzati solo agli scopi per cui sono stati raccolti
- Le terze parti: il trasferimento di dati personali a terzi soggetti è subordinato al consenso del titolare degli stessi
- L'accesso individuale: ciascun individuo ha il diritto di essere a conoscenza dei dati personali in possesso delle diverse organizzazioni
- La correttezza: i dati dovrebbero essere accurati e continuamente aggiornati

In altri paesi come ad esempio negli USA viene adottata una visione più liberale che comporta lo svantaggio in termini di minore protezione dei dati, ciononostante un nucleo di diritti fondamentali viene comunque garantito dalla carta costituzionale. In questo ordinamento non si può impedire alle organizzazioni di fruire dei propri dati

impropriamente, ma viene garantito il diritto ai consumatori di citare in giudizio i responsabili del danno.

SICUREZZA ED ANONIMATO: quando parliamo di sicurezza dei dati non bisogna concentrarsi esclusivamente sulla perdita di denaro dovuta ad eventi come il furto di codici delle carte di credito o alle multe pecuniarie inflitte dalle *authorities* alle organizzazioni, piuttosto l'area più colpita è quella reputazionale e la conseguente perdita di fiducia da parte del consumatore frodato. Tutto ciò al giorno d'oggi avviene con una facilità sempre maggiore, in quanto basta un qualsiasi *device* per poter *hackerare* i dettagli di tutte le carte di credito del Regno Unito. È per questi motivi che le imprese stanno investendo in maniera sempre crescente nell'information security ed è stato stimato che nel 2018 la spesa mondiale per contrastare il furto di informazioni è stata di 93 miliardi di dollari, cifra comunque esigua se confrontata con i 3 trilioni di dollari di danni dovuti ai furti di informazioni (Piero Boccellato, 2017). Le Pmi per poter affrontare questa sfida hanno optato per *l'outsourcing* e il *cloud service* in modo tale da poter contenere i costi pur garantendosi un buon livello di protezione dell'infrastruttura. Volgendo poi l'attenzione sul problema dell'anonimato ciò che viene immediatamente in rilievo è che la regolamentazione in materia di dati si riferisce esclusivamente a quelli personali che possono essere direttamente associati ad un determinato individuo. Di conseguenza la legge sulla privacy troverà applicazione solo nel caso in cui le informazioni di cui un'organizzazione è in possesso permettano l'identificazione di uno specifico individuo. Solo i dati anonimi possono essere venduti e condivisi senza la necessità di chiedere l'autorizzazione. Le organizzazioni per tutelarsi da possibili azioni legali hanno adottato vari stratagemmi, ad esempio piuttosto che condividere il dato preciso "il soggetto ha 37 anni" preferiscono usare la forma "il soggetto ha un'età che oscilla tra i 35 e i 40 anni".

PROCESSI DECISIONALI: l'utilizzo crescente di analisi predittive fa sorgere molteplici questioni su come utilizzare determinati dati nei diversi processi decisionali. Il nocciolo del problema è di riuscire a comprendere quale tipologia di dati appartenenti ad un determinato individuo sia etico utilizzare; per trovare una soluzione è necessario fare un utilizzo coscienzioso dei dati. Ogni persona ha diritto a un trattamento non discriminatorio, in pratica andrebbero tralasciate ai fini commerciali tutte quelle informazioni su cui gli individui non hanno controllo o che comunque non possono

cambiare volontariamente. È ugualmente rilevante sul piano etico quale sia l'impatto che tali decisioni hanno sugli individui. Una visione etica e corretta imporrebbe il tenere conto simultaneamente di tre aspetti:

- Il beneficiario di tali decisioni
- L'impatto che tali decisioni hanno sul beneficiario
- Il grado di mutabilità dei dati da cui si originano i processi decisionali

Quando le organizzazioni vogliono adottare delle decisioni al fine di trarne un beneficio per loro stesse dovranno necessariamente valutare l'impatto che esse possono avere sui destinatari, specialmente se si tratta di decisioni fondate su dati che non possono essere controllati dall'individuo. Al contrario quando il beneficiario della decisione è l'individuo, c'è una minore ponderazione nell'adozione delle decisioni ed una maggiore propensione all'utilizzo di quanti più dati si hanno. (Steven Finlay, 2014)

CAPITOLO SECONDO

BIG DATA PER IL MARKETING

2.1 BIG DATA ADVERTISING

Quello che si chiedono principalmente le imprese, come banche e assicurazioni che rappresentano i pionieri in questo campo e che maggiormente dispongono di dati sui propri clienti, è quale sarà l'effetto che questa gigantesca raccolta e analisi di dati apporterà al valore delle imprese.

Per rispondere a questo quesito ci viene in aiuto uno studio basato su un campione statistico secondo il quale “Nel caso in cui sia già presente una buona struttura di data *analytics* e venga implementata una big data *strategy* nella giusta maniera, l'impresa vedrà crescere la sua performance del 4-5%, a seconda della quantità e della qualità dei dati a sua disposizione. Questo incremento sarà molto maggiore nel caso in cui le imprese non abbiano ancora raccolto dati sui comportamenti dei consumatori, in questo caso la performance può incrementarsi anche fino al 10%. (Steven Finlay)

È ormai chiaro che qualsiasi attività che svolgiamo lascia una traccia/produce dei dati. Quando queste attività si riferiscono a consumi (ricerca di informazione sui prodotti, comparazione di offerte) o a interazioni di mercato (scambi di informazioni ed opinioni sui prodotti) i dati generati sono quelli che le imprese dovrebbero accaparrarsi a tutti i costi, e con cui cercare di costituire il vantaggio in termini di performance che si prefiggono.

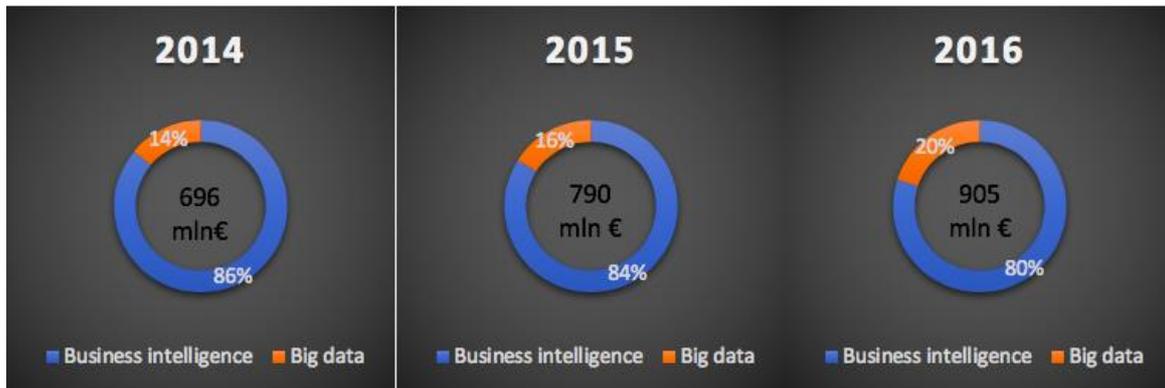


Fig. 31 .

I concetti su cui concentrarsi per capire l’impatto dei *big data* sull’ *advertising* sono 3:

- la possibilità di costruire piattaforme e reti di gestione e condivisione delle informazioni raccolte nella logica della *platform economy*.
- l’obiettivo di ottimizzare e incrementare l’efficienza dei processi per diminuire i costi e aumentare velocità e precisione dei diversi attori.
- L’ incremento del valore creato nelle interazioni e relazioni pubblicitarie.
(Andreina Mandelli)

Le imprese hanno bisogno di capire cosa devono mettere all’ interno del quadro *dell’advertising* e cosa invece debbono lasciare fuori perché fuorviante e privo di valore. A questo scopo soccorrono in aiuto Wedel e Kannan che individuano le principali quattro dimensioni su cui i *big data* insistono e con cui influenzano il *marketing*:

- La profilazione dei consumatori: che si concentra soprattutto nell’ individuazione della *buyer personas*, termine usato nel marketing per identificare e comprendere il cliente tipo nelle sue caratteristiche personali, oltre che socio-demografiche. Di fatto si tratta di un personaggio fittizio delineato nelle sue peculiarità comportamentali e che rivela gli “*insight*” d’acquisto, facendo emergere cosa pensano e cosa fanno i potenziali clienti quando cercano una soluzione a un loro problema o necessità. In tale modo si può confezionare un advertising “*taylor made*”

¹www.osservatori.net, peso e crescita dei big data nel mercato italiano delle analytics

- La possibilità di indirizzare le decisioni riguardanti il marketing operativo (prodotto, prezzo, comunicazione e vendita)
- La possibilità di arrivare ad ogni singolo consumatore per migliorare la precisione dell'azione di marketing a fronte di un aumento dell'eterogeneità dei comportamenti
- La necessità di controllare la sicurezza e protezione della privacy nella relazione con i propri clienti

2.2 INTRODUZIONE AL PROGRAMMATIC ADVERTISING:

Il *programmatic advertising* consiste nell'applicazione dei *big data* all'attività di advertising, permettendo così l'automatizzazione di alcuni processi fondamentali che mettono d'accordo l'offerta (editori) e la domanda di pubblicità (aziende). Non è un segreto che l'industria del marketing digitale ami l'innovazione ed è quindi in uno stato di costante cambiamento e sempre alla ricerca dell'ultima grande novità, ed oggi la grande novità si chiama automatizzazione. Questa pratica consente di acquistare e vendere in maniera automatica spazi o contratti pubblicitari grazie a delle piattaforme che permettono l'incontro anche con aste in tempo reale *real-time bidding* di domanda e offerta. Questi *marketplace* agiscono grazie a dei *software* dotati di intelligenza artificiale che consentono un *matching* sempre aggiornata e complesso che si nutre di enormi flussi dati riguardanti i profili e i comportamenti dei consumatori. Quello che sono in grado di fare questi algoritmi di intelligenza artificiale consiste nella modellizzazione e nella previsione dei comportamenti dei propri *target* grazie ai dati storici e ai flussi di dati in tempo reale. Le caratteristiche principali del principio della pubblicità programmatica sono (Busch 2016):

- Granularità: cioè piena conoscenza delle opportunità di *impression* dell'annuncio insieme ai propri parametri generali quali destinatari specifici e ambiente pubblicitario di riferimento.
- Trading in tempo reale: consiste nel poter prendere decisioni in tempo reale riguardo una specifica pubblicità o una possibilità di *impression*

- Informazioni in tempo reale: consentono di valutare la reale portata dell'utilità del *advertisement* sulla base delle sue caratteristiche specifiche e dei dati empirici pertinenti raccolti fino a quel momento.
- Creazione in tempo reale: gli inserzionisti pubblicano un annuncio (ove possibile dinamico e basato sui dati) che meglio si adatta alla situazione tenendo conto di tutti i dati a propria disposizione immediatamente dopo essersi aggiudicati lo spazio pubblicitario.
- Automazione: consiste in un processo di prenotazione e pubblicazione degli *advertisement* in maniera completamente automatizzata.

In breve: "La pubblicità programmatica descrive il processo di pubblicazione automatizzato di annunci digitali in tempo reale basati sulle opportunità di *impression* dei singoli annunci" (trad. Bardowicks e Busch2013).

2.3 COME FUNZIONA IL PROGRAMMATIC ADVERTISING:

Facebook e Google sono stati i primi ad adottare e a promuovere sistemi che forniscono a milioni di clienti la possibilità di gestire le loro campagne pubblicitarie automatizzate e guidate da dati in tempo reale. I fornitori di tecnologie di *programmatic advertising* hanno adottato le principali caratteristiche precedentemente spiegate, vale a dire il meccanismo dell'asta, la granularità e la trasparenza (Heimann2013). Questo non solo richiede la formazione e il "confezionamento" di pacchetti completi, ma soprattutto concentra l'offerta globale di media online all'interno di una sola grande piattaforma. Questa aggregazione riduce drasticamente il numero di intermediari da connettere all'interno della suddetta piattaforma di *trading* che agisce in tempo reale, rendendo così possibile il *trading* automatizzato. Perché ciò accada, gli attori del mercato hanno bisogno di *standard* globali dettati da protocolli, formati e metodi da cui possano attingere in modo da poter incrementare il potenziale di crescita del mercato.



Figura 42

In pratica gli algoritmi di intelligenza artificiale riescono a modellizzare e prevedere i comportamenti dei consumatori partendo dai dati storici e da quelli che affluiscono in tempo reale. Il processo avviene grazie a tre diverse piattaforme:

- DSP (*demand side platform*) hanno l'obiettivo di facilitare l'acquisto dell'inventario pubblicitario sul mercato aperto, dando la possibilità di raggiungere il *target* profilato continuamente dalle DMP.
- SSP (*supply side platform*) tramite cui gli editori possono fornire la propria offerta, in pratica si occupa di distribuire la pubblicità sui vari siti degli editori.
- DMP (*data management platform*): si occupa di profilare continuamente i soggetti in modo da identificare le caratteristiche e le *buyer personas* tramite *digital cookies*³ e dati da *mobile app*.

² Andreina Mandelli, 2016, big data marketing, le piattaforme e i processi integrativi nel programmatic advertising viene spiegato in che modo l'AI riesce a confezionare in tempo reale grazie ai cookies delle specifiche offerte per i consumatori

³ "I cookies sono piccoli file che vengono memorizzati sui dispositivi utilizzati per navigare su internet, siano essi computer, cellulari o tablet. Quando un utente naviga su un sito web, solitamente scarica dei cookies presenti su quel sito. Questi piccoli pezzi di codice svolgono certamente delle funzioni importanti ed utili per lo stesso visitatore, ma ve ne sono alcuni più intrusivi che tengono traccia della navigazione

Le tre piattaforme funzionano continuamente e in tempo reale. Ed è questo il motivo per cui il *programmatic advertising* è anche chiamato *real time buying*. Grazie a queste piattaforme può avvenire l'acquisto delle *unit* pubblicitarie che permettono l'incontro tra *publisher* (che vendono le *impression*) e *advertiser* (che cercano spazi per trasmettere il messaggio alla propria platea). La conseguenza significativa di questo processo automatizzato guidato dai dati è una migliore efficienza e una maggiore immediatezza.

Le aziende ed i *Web marketer* possono avere accesso ad un'enorme quantità di *media* e *impression* oppure possono optare per un vero e proprio scambio privato con cui i *publisher* possono controllare l'accesso dei *buyer* e le relative offerte. Altre modalità di acquisto sono il sistema *open auction* (o asta libera), in cui l'asta è aperta a molteplici inserzionisti pubblicitari, e il sistema *guaranteed direct*, che consiste nella vendita automatizzata di spazi in giacenza tra *buyer* e *seller* attraverso una piattaforma che gestisce in maniera diretta l'acquisto e l'offerta.

Negli USA nel 2017 il 78 per cento di tutto *l'advertising display* è avvenuto sotto la forma di *programmatic advertising* (eMarketer). In Italia siamo ancora molto lontani da questi numeri anche se il mercato dimostra potenziale di crescita. Infatti stando a quanto riportato dall'osservatorio *internet media* del politecnico di Milano c'è stato un ragguardevole aumento degli investimenti in *programmatic advertising* che nel 2015 sono aumentati del 113 per cento rispetto all'anno precedente facendo schizzare il giro d'affari a 234 milioni di euro, per rendere l'idea questa cifra nel 2012 corrispondeva ad appena 5 milioni di euro. Alla fine del 2016 questo settore ha visto un aumento degli investimenti di oltre 300 milioni di euro, arrivando ad avere una quota di mercato tra il 25% e il 30% del totale del *display advertising* (Osservatorio Internet Media Politecnico di Milano).

2.4 FACEBOOK PIONIERE DEL PROGRAMMATIC ADVERTISING

Fin dal suo lancio, la pubblicità su Facebook ha assunto varie forme rivoluzionarie ed evolutive, cercando quanto più possibile di fare attenzione a bilanciare gli interessi dei *marketer* con quelli degli utenti che non vogliono essere sovraccaricati di pubblicità. I *social network* hanno possibilità di offrire una pubblicità mirata che insieme alle opzioni

di *targeting* disponibili possono creare una esperienza omnicanale⁴ per l'utente finale. Facebook fornisce non solo la propria piattaforma ma anche agenzie pubblicitarie che pianificano ed eseguono campagne per conto di inserzionisti e aziende, oltre a mettere a disposizione i propri sviluppatori di tecnologia pubblicitaria che svolgono il ruolo di intermediazione tra gli inserzionisti e Facebook. Quando Facebook ha aperto le porte alla sua API Ads (*Application Programming Interface*) nel 2011, sono stati sviluppati una serie di strumenti per aiutare gli inserzionisti a gestire le campagne di annunci social. La creazione dell'API di Facebook ha consentito lo sviluppo di funzionalità avanzate e ha consentito agli inserzionisti di lavorare in modo più efficiente. La crescita di questo ecosistema ha consentito un ciclo di *feedback* continuo, contribuendo al successo degli inserzionisti dei *social network*. Con strumenti sofisticati che consentono strategie avanzate e test granulari, l'API di Facebook ora offre un approccio più intelligente alla pubblicità sui *social network*. Ciò ha portato risultati di *business* tangibili agli esperti di *marketing*: ad esempio P&G ha visto una riduzione dell'86% del *cost per engagements* dopo l'adozione di un *tool* di gestione degli annunci social. (Dawson & Lamb, 2015)

2.5 CONSUMER-CENTRIC PROGRAMMATIC ADVERTISING

2.5.1 L'EVOLUZIONE DEL PROGRAMMATIC ADVERTISING

La pubblicità programmatica si è evoluta dal semplice *retargeting* e *audience buying*⁶, fino ad arrivare al più moderno *consumer centric marketing*. L'utilizzo sempre maggiore di questo metodo, unito ad una maggiore sofisticatezza ha portato a una ridefinizione dei modelli di business dei fornitori di *programmatic advertising* che spaziano dal *packaging* del prodotto, a servizi completamente personalizzati. Gli inserzionisti hanno sempre più possibilità di scegliere la configurazione ideale per la proposta pubblicitaria, che sposta

⁴Per Omnicanalità si intende la capacità di seguire il comportamento dei clienti, di saper gestire in tempo reale la comunicazione su diversi punti di contatto (online ed offline) e di adottare strategie di marketing e CRM basate sul dato (data-driven) come elemento guida". (e-business consulting)

⁵ "Con l'offerta cost per engagement (CPE), gli inserzionisti pagano solo quando gli utenti interagiscono attivamente con gli annunci"

⁶ è una modalità di acquisto del digital advertising che consente all'inserzionista di selezionare esclusivamente i contatti che corrispondono al target desiderato.

il fulcro della discussione dal "meglio fatto in casa o fatto in agenzie di *marketing*" a "qual è la giusta combinazione di fatto in casa e in *outsourcing* per le nostre esigenze".

Il *marketing* di oggi richiede che i professionisti del *marketing* si concentrino sull'intero sistema di comunicazione e su come ottimizzarlo per cambiare i risultati della propria attività. Viviamo nell'era del contenuto e delle connessioni dove dobbiamo capire come canali e contenuti possono lavorare al meglio insieme per aumentare la *performance* della comunicazione. L'obiettivo è avere una strategia di distribuzione chiaramente definita, garantendo che non ci siano vicoli ciechi o sprechi derivanti da contenuti sconnessi tra loro. La pubblicità programmatica è un enorme passo avanti per consentire di fornire contenuti adeguati a ogni singolo consumatore per guidare la *customer relationship* in maniera più efficace. La combinazione di tecnologia in tempo reale con una comprensione sempre più olistica del profilo utente individuale consente di mantenere una comunicazione di tipo *one-to-one*.

Il *programmatic advertising* ha fatto molta strada partendo da una versione 1.0 per arrivare ad oggi con l'ultima versione 3.0.



Figura 57

⁷ Programmatic advertising: fonte <https://link.springer.com/content/pdf/10.1007%2F978-3-319-25023-6.pdf>

- *Re-targeting*: è uno strumento molto potente per coinvolgere nuovamente i consumatori che hanno manifestato un interesse ma non si sono convertiti al prodotto. I dati di tipo *targeting* sui profili dei consumatori vengono raccolti attraverso una procedura molto semplice che consiste nel posizionamento di *tracking pixel* sul proprio sito *Web* lungo il percorso di conversione, in genere concentrandosi su pochi tipi di dati (tuttavia spesso fornendo migliaia di offerte personalizzate basate sull'ultimo prodotto che il consumatore ha visto). Le offerte in tempo reale di *ad exchange* negoziate nell'ecosistema del *programmatic buying and selling* in maniera simile alle piattaforme di trading utilizzate nei mercati borsistici, sono un modo molto efficiente per acquistare una grande copertura anche se la qualità del posizionamento occupa una posizione secondaria rispetto alla possibilità di trovare il maggior numero di profili utente convertibili. Il *re-targeting* ha un vantaggio di costo notevole, ma ha una portata limitata nel *media mix*.
- *Audience buying*: è una modalità di acquisto del *digital advertising* che consente all'inserzionista di selezionare esclusivamente i contatti che corrispondono al target desiderato. Mentre il *retargeting* si concentra sulla fine del processo di acquisto, l'*Audience Buying* è maggiormente utilizzato per raggiungere i profili dei consumatori che non sono ancora stati coinvolti dall'inserzionista. Società tecnologiche come Quantcast, XAXIS o Rocketfuel raccolgono una vasta quantità di dati anonimi dei consumatori per costruire i profili *target*. Tramite l'*Audience Buying* si cerca di raggiungere i giusti gruppi di consumatori piuttosto che acquistare annunci in un particolare segmento editoriale.
- *Consumer centric* lungo il *funnels* d'acquisto: Il grande ritorno sugli investimenti ottenuti dal *retargeting* e dall'*audience Buying* ha convinto molti inserzionisti sul

⁸ Con il termine Marketing funnel si identifica quel percorso ad imbuto formato dai diversi step compiuti dagli utenti prima di trasformarsi in clienti che convertono. Più il funnel è corto meno possibilità ci sono che l'utente si disperda.

valore che il *programmatic advertising* può avere nel *sales process* e sul valore aggiunto dai dati e dalla maggiore qualità del posizionamento. Nel *programmatic advertising 3.0*, gli inserzionisti passano ad una visione maggiormente incentrata sul consumatore, essi infatti combinano il maggior numero possibile di informazioni in un profilo utente d'insieme, usando quanti più dati possibile combinando ad una segmentazione sofisticata dei visitatori del sito *Web* dati di vendita e arricchendoli con dati di terze parti. Tali profili vengono spesso applicati a un processo di acquisto molto più selettivo, mirando a preservare la qualità degli *asset* strategici (marchio, notorietà) e una selezione degli inserzionisti maggiormente qualitativa. Ciò può essere facilitato tramite le *Private Marketplace Deals* (PMD) che combinano i benefici della negoziazione diretta (maggiore controllo su dove viene promossa la campagna pubblicitaria, accesso a formati di annunci premium e prezzi negoziati) con i vantaggi della raccolta delle *impressions* basata sui dati in tempo reale.

2.5.2 CONSUMER CENTRIC MARKETING: UNA VISIONE DEL SINGOLO CONSUMATORE

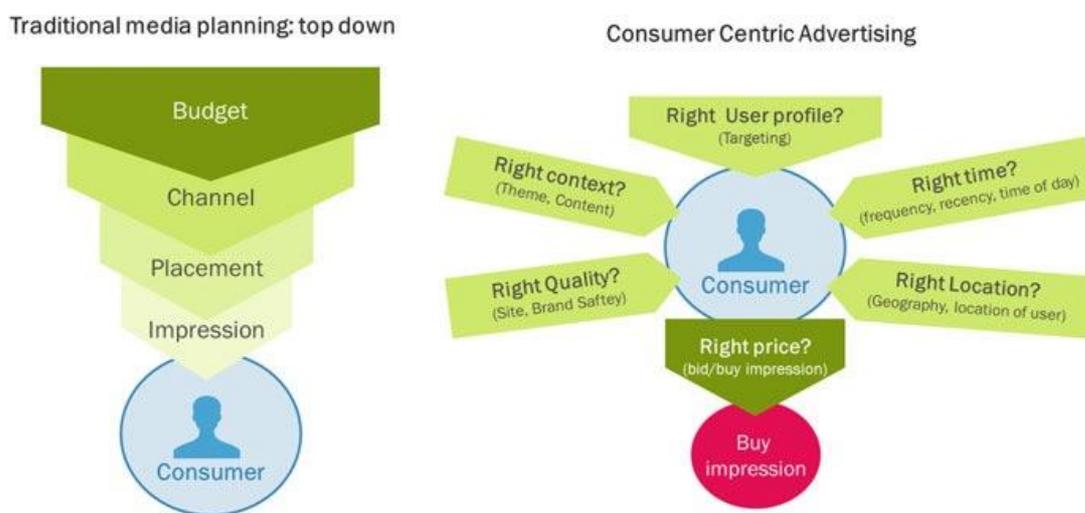


Figura 69 *Traditional media planning vs consumer centric*

La pubblicità programmatica incentrata sul consumatore è un cambiamento radicale nell'approccio alla pianificazione dei mezzi di comunicazione. Tradizionalmente, i *media*

⁹ <https://link.springer.com/content/pdf/10.1007%2F978-3-319-25023-6.pdf>

planner assegnavano il proprio *budget* attraverso canali intermedi all'interno di spazi specifici (in base a affinità, raggiungimento di gruppi di clienti o rendimenti storici) per arrivare a suscitare delle *impression* che si sperava potessero raggiungere i giusti consumatori. Il nuovo modello di pubblicità incentrata sui consumatori (a destra nella figura) avviene in due fasi:

- 1 I profili dei consumatori sono individuati utilizzando tutti i dati disponibili, combinando dati provenienti da prime, seconde e terze parti.
- 2 Quando il DSP individua il profilo di un cliente, i processi complessi valutano dove quel singolo profilo si trova nel percorso di acquisto, qual è il valore di quel profilo e quindi quale sarebbe l'*impression* giusta da fornire in quel momento. Nel fare tale valutazione, i dati del profilo del consumatore vengono incrociati con dati aggiuntivi per rispondere alle seguenti domande: (a) giusto contesto? (b) giusta qualità? (c) momento giusto? (d) posizione corretta? (e) prezzo giusto?

Tutti i punti sopra indicati, combinati con un modello di previsione del ROI determinano il valore della pubblicazione di un annuncio di questo tipo su quello specifico sito, in questo determinato momento, portando così a formare il prezzo dell'offerta. Questo nuovo modello *Consumer Centric* richiede spesso una riorganizzazione significativa del processo di raccolta dei dati digitali per combinare tutti i dati che possono trovarsi nei *database* in questo momento. A differenza delle campagne a breve termine per prodotti o di offerte specifiche, la pubblicità di tipo *Consumer Centric* è sempre attiva, il prodotto e le offerte sono selezionati sulla base degli interessi dei consumatori piuttosto che sulle decisioni di marketing dall'alto verso il basso. Questo porta a una diversa strategia basata su una costante ridefinizione della campagna di marketing *core*, al contrario nelle campagne più tradizionali di comunicazione di tipo push vengono aggiunte alla campagna di base sempre attiva le pratiche di *agenda setting* e lo *storytelling*. (Oliver Gertz e Deirdre McGlashan 2016)

2.6 RISCONTRI, RAZIONALITÀ E VALORE

Prevalentemente sino ad oggi l'obiettivo degli attori coinvolti (agenzie, editori e data broker) era la ricerca dell'ottimizzazione dei processi. Infatti le decisioni di comunicazione avvengono con indirizzo di alta precisione grazie alla micro segmentazione ed anche alla possibilità di raggiungere il consumatore con un messaggio *taylor made* specifico e contestualizzato. (Mandelli)

Nonostante infatti la creazione di piattaforme di business che dovrebbero favorire la condivisione dei *data stream* ed i servizi di analitica *AI-based* che teoricamente dovrebbero favorire l'estrazione condivisa di valore, il trend che le economie digitali stanno seguendo sembra avere come riferimento un meccanismo basato sulla gerarchia e sulla dominanza del mercato. Questa è proprio la logica che negli ultimi vent'anni ha portato alla concentrazione economica a cui siamo arrivati oggi (Facebook, Google, Amazon). La logica del valore invece che sta cercando di diventare centrale con le nuove filosofie di *marketing customer-centric* è spinta dalle nuove tecnologie e dai nuovi ambienti digitali collaborativi. In questa ottica il ruolo dei consumatori non è più quello di un soggetto passivo senza ruolo ma di attivo condivisore di risorse ed informazioni sulle piattaforme.

Ovviamente il focus sull'ottimizzazione resta di fondamentale importanza perché consente una sempre più precisa e specifica offerta in grado di creare valore anche per il cliente, quando essa genera benefici tangibili al netto dei costi richiesti delle interazioni. La maggior efficienza di spesa per l'impresa (ottimizzazione) può anche sposarsi con la creazione di valore per il cliente, infatti quanto più grazie all'utilizzo informazioni precise e dettagliate (big data) si può aumentare la contestualizzazione del messaggio pubblicitario tanto più aumenterà l'utilità per il consumatore che percepirà una *brand awareness* migliore.

In questo complicato passaggio le imprese devono stare attente al fenomeno del *ad blocking* dovuto al fatto che molti consumatori si sentono infastiditi e bloccano *l'advertising* sui proprio *device*. Sono circa 309 milioni le persone aderiscono a questa pratica, se la cifra non sembra molto alta è solo perché in molti non sanno che possibile bloccare le pubblicità (Mobile adblocking Report 2016). È necessario che si investa

seriamente in policy di protezione che non facciano sentire limitazioni della libertà del consumatore.

In conclusione nei prossimi 5-10 anni il principio programmatico sarà alla base tutte le comunicazioni con i consumatori poiché l'adozione di pubblicità programmatica e la velocità di apprendimento da parte degli attori del mercato porteranno a spostamenti di quote di mercato da parte dei media e degli inserzionisti. (Busch)

CAPITOLO TERZO

3 DIGITAL PLATFORM E BIG DATA

3.1 I BIG DATA GUIDANO LA RIVOLUZIONE DEI BUSINESS

Nella moderna economia digitale, la struttura del mercato è caratterizzata dalla presenza di grandi operatori che si sono guadagnati nel tempo la loro posizione e questa concentrazione ha stimolato un dibattito sul rischio associato al potere di mercato. Anche molti opinionisti influenti sulla stampa mondiale chiedono un intervento delle autorità antitrust, altri invece sono del parere contrario.

Ci concentriamo quindi proprio sul business delle cosiddette *digital platforms*, che sono i principali attori nei mercati digitali e suscitano preoccupazione a causa del loro enorme potere di mercato (Brousseau e Pénard, 2007; Smith e Telang, 2016). Il vantaggio competitivo di queste aziende si basa sui dati che possono raccogliere dai loro utenti e clienti al fine di alimentare modelli predittivi in grado di determinare probabilisticamente le preferenze e i comportamenti di acquisto. I flussi massicci ed eterogenei di dati generati da qualsiasi tipo di transazione Internet o impronta digitale sono la risorsa fondamentale per migliorare la business intelligence incentrata sul cliente e anticipare ciò che i consumatori desiderano. Internet è stato un fattore chiave nella generazione e raccolta di enormi quantità di informazioni, sia *offline* che *online* anche se bisogna prestare una grande attenzione quando si selezionano le informazioni su cui poi basare il proprio business; per parafrasare Herbert Simon, “una ricchezza di informazioni significa una carenza di attenzione da parte dei suoi destinatari, poiché l'attenzione dovrebbe essere allocata saggiamente data la sovrabbondanza di stimoli informativi” (Simon, 1971).

L'ascesa dei mercati digitali basati sui dati ha come caratteristica principale quella di favorire una forma di concentrazione del settore (Nuccio, Guerzoni 2018). I principali *driver* di questo risvolto sono due, da un lato i consumatori che sopraffatti dall' enorme massa di informazioni non riescono ad avere una visione d'insieme del mercato e delle opportunità disponibili, mentre dall' altro lato le imprese che desiderano trarre profitti

rilevanti devono aumentare il più possibile la propria visibilità nei confronti dei consumatori. A questo punto intervengono gli intermediari digitali che aggregano l'offerta filtrandola a supporto dei consumatori, al contempo i consumatori sceglieranno le aziende maggiormente visibili sulle piattaforme, favorendo così una forma di concentrazione del settore. Per riassumere, le piattaforme condividono tra loro vantaggi competitivi simili. In primo luogo, coordinando gli agenti, riducono la quantità di attenzione necessaria ai consumatori per navigare nella grande quantità di informazioni offerte dall'economia digitale. Per questo motivo, una caratteristica distintiva delle piattaforme è la ricerca di qualcosa nell'infinita gamma di informazioni fornite da Internet, siano esse notizie, video, automobili, appartamenti o hotel. In secondo luogo, le stesse aziende condividono anche competenze specifiche e una struttura fisica per l'analisi dei big data generati dal cliente al fine di ottimizzare abbinamento, marketing e logistica, a seconda del tipo di servizio offerto. “La concentrazione risultante è sorprendente. Le cinque più grandi *public companies* mondiali di tecnologia, ovvero Apple, Alphabet (Google), Microsoft, Amazon e Facebook, valgono collettivamente quasi 4,2 trilioni di dollari. Nel 2008, Google ha lanciato Chrome, che 10 anni dopo controllava circa il 70-80% del mercato mondiale dei motori di ricerca seguito da Firefox e Internet Explorer con una quota di mercato del 13% ciascuno, inoltre, se guardiamo a mercati specifici, i servizi basati su *Internet* stanno prendendo quote di mercato da servizi tradizionali equivalenti e tendono ad essere molto concentrati. Ad esempio, il mercato delle agenzie di viaggio online (OTA) è fondamentalmente un duopolio di Expedia e Priceline (booking.com), che detengono i due terzi del settore globale delle prenotazioni di viaggi online e il 90% di quello degli Stati Uniti” (Nuccio, Guerzoni).

Non capire questi nuovi modelli di business è realmente pericoloso per le imprese (Van Alstyne 2016), questi mercati e settori evolveranno le loro logiche non solo nei loro servizi, ma anche nelle loro risorse grazie all'accesso a nuove forme e volumi di dati, chi rimarrà indietro perderà la propria capacità di offrire un vantaggio competitivo.

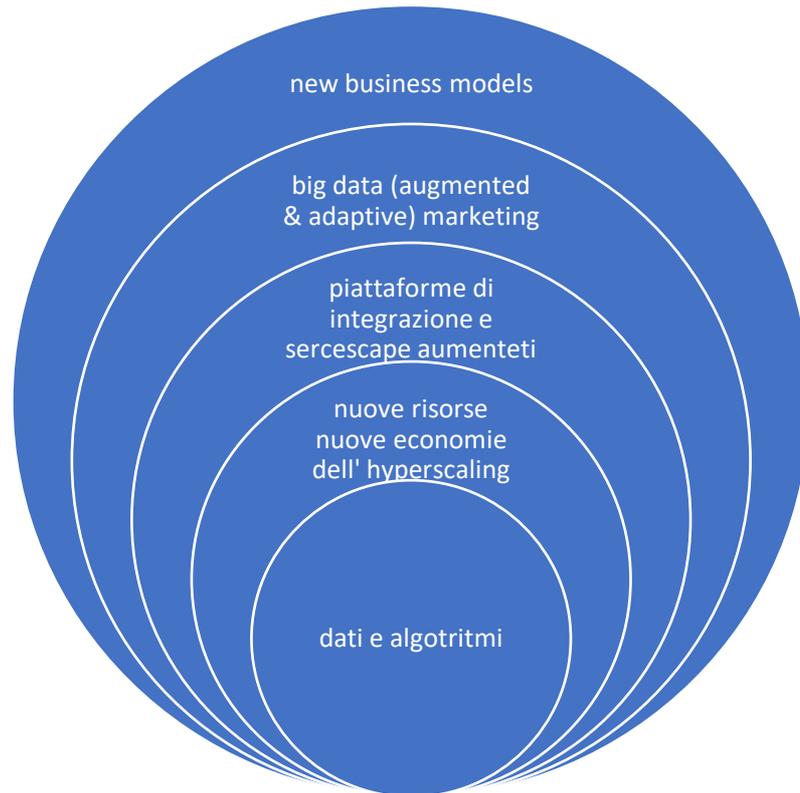


Figura 51

3.2 IN CHE MODO I BIG DATA ALIMENTANO LA SHARING ECONOMY

Oggigiorno si parla moltissimo della *sharing economy*, fenomeno in costante espansione che vede tra le sue fila aziende come Uber, Airbnb, Enjoy e BlaBlaCa, le quali sono entrate prepotentemente nella nostra vita rivoluzionando il vecchio modello socio economico capitalistico. Non si può parlare però di *sharing economy* senza risalire a quelle che sono le basi per la sua esistenza, e cioè la *platform economy*, ma soprattutto il pane quotidiano di questo fenomeno che sono senza dubbio i *big data*.

Ma prima di capire come i *big data* alimentano questa rivoluzione è importante dare qualche numero per rendersi realmente conto della portata di questo fenomeno e del cambiamento in atto.

Un team dell'Osservatorio Startup Intelligence del Politecnico di Milano si è posto l'obiettivo di inquadrare e definire il fenomeno della *sharing economy*. “La *Sharing*

¹ Grafico, Mandelli A. 2017.

Economy è da anni un tema caldissimo: una semplice ricerca su Google dà oltre 18 milioni di risultati. E muove enormi volumi di denaro: molte startup sono diventate *unicorn2* – Uber, Didi, Airbnb, Lyft, BlaBlaCar – e in tutto abbiamo rilevato investimenti di oltre 4 miliardi di dollari in 195 startup, di cui 26 italiane che hanno raccolto 23 milioni di dollari” (Ghezzi 2018). Delle italiane, la più finanziata è Supermercato24 (oltre 5 milioni), seguita da ProntoPro (servizi professionali occasionali) con 3,7 milioni, e Moovenda (consegna di cibo a domicilio) con 2,2 milioni.

Passiamo quindi ad una lettura dell’aspetto che più ci interessa di questo fenomeno, ossia individuare in che modo i *big data* fungono da driver per la *sharing economy*.

- Possono aiutare a soddisfare le esigenze

I consumatori esprimono ciò a cui sono interessati lasciando tracce sui dispositivi come abbiamo precedentemente spiegato, e proprio tramite questi dati che sono trasmessi in tutto il globo che successivamente il fornitore di servizi può attivarsi per soddisfare questa domanda. Su ampia scala, gli algoritmi di *big data* sono in grado di connettere le persone perfino riguardo ad aspetti molto specifici dei prodotti/servizi che stanno cercando. Questi algoritmi utilizzano dati come posizioni GPS e altri dati personalizzati per connettere i consumatori con i fornitori che possono soddisfare maggiormente le loro esigenze. In effetti, è questa efficienza che rende così popolari servizi come Uber. I consumatori apprezzano il fatto di poter soddisfare le proprie esigenze più specifiche soltanto con un click sui loro telefoni.

- Concedono maggiori opportunità alle *startups*

I *big data* sono anche il modo in cui le start-up possono entrare nel mercato in modo economico. I dati offrono alle *startup* l’opportunità di interpretare e anticipare facilmente le esigenze dei consumatori, e sono successivamente in grado di informarle sulla qualità percepita del servizio offerto e consentono persino alle *startup* di ricevere un ampio e diffuso supporto dalle comunità di *crowdfunding*. I siti di *crowdfunding* sono elementi chiave nell’economia della condivisione, in cui i cittadini privati possono investire nelle

² Startup che valgono più di un miliardo di dollari.

idee degli sviluppatori e vederle diventare realtà. I siti di finanziamento *crowdfunding* sono stati utilizzati per finanziare qualsiasi cosa, dai film alle imprese, ma quando si tratta di servizi, le potenziali aziende possono utilizzare i dati per mostrare scientificamente ciò che i potenziali investitori stanno cercando in un investimento. Possono quindi utilizzare tali dati per garantire che la loro idea goda di un'attrattività maggiore e quindi ricevere più finanziamenti.

- Danno la possibilità agli imprenditori di far crescere le piccole imprese

Una cosa che si può notare di molte delle società che partecipano alla *sharing economy* è che in realtà esse non forniscono alcun tipo servizio da sole. Ciò che invece forniscono è una piattaforma. Attraverso queste piattaforme, conducenti, scrittori, proprietari di immobili e rivenditori sono in grado di far crescere senza fatica la propria attività semplicemente pubblicando le proprie informazioni sulle *digital platforms* e facendole vedere ai consumatori che necessitano di prodotti/servizi.

Di conseguenza, ci sono molti professionisti che hanno una completa autonomia sulla loro carriera e un facile accesso a clienti e clienti che hanno bisogno esattamente di ciò che questi professionisti possono offrirgli.

- Sono utili anche per i consumatori

I clienti che usufruiscono dei servizi della *sharing economy* spesso li preferiscono a quelli delle imprese tradizionali. In effetti, il 60% dei viaggiatori che usano AirBnB lo preferiscono alla tradizionale esperienza alberghiera. Solitamente il motivo è che siti come Uber e AirBnB mantengono i prezzi competitivi. I clienti ricevono grazie a queste piattaforme molte offerte diverse di prezzo e sono quindi certi di scegliere l'opzione più conveniente disponibile.

- Assicurano la sopravvivenza del migliore

Sia che si stia acquistando un bene online, assumendo uno scrittore per lavorare al proprio blog o semplicemente prenotando un taxi tramite un'app, ci si troverà a lavorare con un completo sconosciuto. Uno dei ruoli più importanti che i big data svolgono nella *sharing*

economy è che consentono ai clienti di leggere recensioni su un *freelancer* prima di assumerlo.

Siti come *Upwork*, *Uber* e *AirBnB* forniscono tutte le recensioni sui loro lavoratori *freelancer* e sui servizi in modo da poter fornire alcune informazioni di base sulla loro storia lavorativa. Questi dati consentono ai migliori lavoratori di eccellere e mettono in contatto i consumatori con i servizi della massima qualità possibile.

3.3 IL CASO AIRBNB

Per capire la reale importanza dei dati per i leader dei *marketplace online*, occorre soffermarsi attentamente sul caso *Airbnb*. La società nasce nel 2008 a San Francisco seguendo la più classica formula delle piattaforme di *sharing economy* connettendo tra loro persone che vogliono fornire dei servizi ad altre che ne vogliono usufruire. Il modello di *business* che tutti conosciamo è alimentato da una *fee* che riceve Airbnb sia dall' *host* che dai *guest* che stipulano un contratto di locazione tramite la piattaforma.

È proprio la piattaforma di *Airbnb* a mettere a disposizione in rete tutti i dati relativi alla disponibilità di appartamenti e proprietà messi in affitto. I dati che vengono via via registrati, integrati con altri provenienti dall'esterno vengono utilizzati non solo per avere maggiori informazioni sui propri clienti, ma anche per supportare le decisioni di *business*, (che hanno portato ad esempio ad ampliare significativamente il *bundle* dei servizi rispetto a quelli previsti in origine). Gli *insight* acquisiti dall' analisi dei *big data* consentono inoltre ad *Airbnb* di concentrare i propri sforzi sull' acquisizione di nuovi *host* presso le destinazioni più richieste. Grazie agli *insight* inoltre la multinazionale consiglia ai propri *host* i prezzi da praticare in linea con quelli che sono suggeriti dalle informazioni in loro possesso (relazione *pricing*-probabilità affitto, caratteristiche dell'immobile, periodo in questione...) potendo sfruttare appunto una grande massa di dati. La crescita sbalorditiva di Airbnb è dovuta al fatto che essa elabora 1,5 petabyte di dati, tra l'altro ospitati sul *cloud* di un'altra *digital platform* (Amazon).

Questo caso dimostra come nella *platform economy* l'impatto trasformativo dei big data ha avuto un effetto "disruptive". È proprio questo l'ambito in cui i *big data* hanno avuto un impatto maggiore, sia sul *marketing* che sul *business* stesso, imponendo un

ripensamento delle modalità di costruzione e progettazione dell'esperienza del cliente (*consumer decision journey*). È avvenuto quindi lo *shift* organizzativo verso la cosiddetta "*sharing economy*" in cui si è passati dal controllo delle risorse al coordinamento delle stesse cercando di valorizzare quanto più l'intero ecosistema.

3.4 BARRIERE ALL' INGRESSO NELLA PLATFORM ECONOMY

In questo paragrafo l'attenzione sarà rivolta a chiarire come un concetto così cruciale per l'economia, ossia quello delle *entry barriers* sia cambiato con l'avvento dei *Big data*. Sebbene si possano trovare in letteratura svariate definizioni di barriere all'entrata, quella proposta da J.S. Bain in *Barriers to new competition* (1956) rimane fra le maggiormente apprezzate: "è il vantaggio di un'impresa che già opera in un mercato rispetto a un qualsiasi potenziale entrante, che si riflette nella capacità della prima di mantenere il proprio prezzo costantemente al di sopra dei livelli che si avrebbero in concorrenza, senza tuttavia attrarre nuove imprese in quel mercato". Secondo la teoria economica, inoltre, un operatore già presente sul mercato non può sfruttare la propria posizione a scapito del consumatore se non è in grado di costruire barriere all'ingresso durevoli, dovendosi poi quindi trovare di fronte al rischio costante di dover competere con potenziali concorrenti (Baumol, 1986). Vediamo adesso come concetti classici dell'economia siano cambiati con l'avvento dei big data.

- I big data essendo beni pubblici riducono le barriere all'ingresso

Per loro natura, i beni di informazione sono concepiti come beni pubblici, anche se le imprese possono provare in maniera artificiale ad escludere alcuni soggetti dall'accaparrarsene. L'uso dei dati da parte di una società non impedisce a un'altra società di raccogliere e utilizzare gli stessi dati. Haucap e Heimeshoff hanno già chiesto una maggiore trasparenza nei mercati, come ad esempio quello dei motori di ricerca dove appunto questi dati sono alla base del corretto funzionamento, ma hanno invece suggerito di non condividere l'algoritmo (che è la fonte del vantaggio competitivo) con possibili concorrenti.

- Velocità: rendimenti sempre meno legati alla storicità delle imprese e valori aggiornati in tempo reale riducono le barriere all'ingresso

La vita dei dati è breve e avere una relazione a lungo termine con i clienti non garantisce il successo in futuro. Sebbene le analisi delle serie storiche dei dati comportamentali possano prevedere le preferenze dei clienti, il loro ritorno su scala diminuisce nel tempo. In effetti, in molte aziende, i modelli predittivi si sono progressivamente spostati verso le tecnologie di *nowcasting*, che analizzano i dati in tempo reale (Familiar and Barnes, 2017). Ad esempio, il modello di business dell'offerta in tempo reale è ormai il modello dominante, come precedentemente esposto, per i mercati della pubblicità *online* poiché può identificare in tempo reale le caratteristiche e gli interessi del pubblico target per ogni *impression* e automaticamente offrire la migliore pubblicità a prezzi ottimizzati. A questo proposito, i *big data* rappresentano non solo un vantaggio significativo per le piattaforme digitali, ma anche per le società non digitali (Lerner 2014). Per le società tradizionali, come quelle nel settore assicurativo e bancario, come i grandi e piccoli rivenditori, o gli operatori logistici e le pubbliche amministrazioni hanno adottato vari dispositivi (ad esempio applicazioni, carte fedeltà o tessere associative) per raccogliere e analizzare informazioni sul comportamento dei propri clienti (Lerner 2014). Ad esempio, uno *smart grids* può misurare come e quando i consumatori consumano più energia e aiutare sia gli utenti che le società di servizi a gestire e pianificare meglio il consumo di energia.

- Volume e varietà: l'accumulazione e l'onniscienza dei dati non rappresentano una barriera d'ingresso

Sebbene il volume sia la principale caratteristica dei *big data*, il loro valore è piuttosto collegato ai contenuti eterogenei che includono dati strutturati e non strutturati, raccolti appositamente e casualmente (Rubinfeld e Gal, 2017). L'accumulo di dati non è di per sé considerato come una barriera all'ingresso, purché non sia impedito ai competitors di fare lo stesso. Non vi è dubbio che le tecniche di *machine learning* e intelligenza

³ secondo quella che è la definizione dell'enciclopedia Treccani: "Sistema elettrico in cui le relazioni di connessione tra i nodi attivi sono integrate e coordinate in modo cosiddetto intelligente, attraverso sistemi telematici automatizzati, per consentire una gestione dell'energia elettrica che risponda ai criteri di efficienza, sostenibilità e sicurezza"

artificiale richiedono l'uso di una grande quantità di informazioni per addestrare i loro algoritmi, ma, senza la selezione di dati dedicati a un problema specifico, il rischio di correlazione spuria, di *overfitting*⁴ e di altri errori predittivi è ancora sostanziale.

- Economie di scala ed esternalità di rete non per forza rappresentano barriere

Le piattaforme sono in genere facili da costruire ma risulta difficile ridimensionarle ogni volta e richiedono una continua messa a punto per essere affidabili ed efficienti. Un grande vantaggio di un approccio funzionante tramite big data è il cosiddetto "effetto di rinforzo", in grado di generare esternalità di rete dirette e indirette. Le esternalità dirette dipendono dalle dimensioni della rete e sono diventate molto importanti per piattaforme come *Facebook*, dove gli utenti conferiscono più valore quanti più "amici" usano lo stesso *social network*. Gli effetti indiretti sono particolarmente efficaci sulle così dette *two-side platforms* come *eBay*, poiché un numero maggiore di utenti da un lato (clienti finali) possono attrarre un numero maggiore di utenti dall'altro (inserzionisti o commercianti).

I vantaggi di scala però non si traducono necessariamente in vantaggi in termini di attrattività per gli inserzionisti. I clienti sono spesso disturbati da un sovraccarico di pubblicità e gli inserzionisti corrono un rischio considerevole scegliendo di mettere tutto il proprio *budget* di *marketing* su un'unica piattaforma in quanto ciò potrebbe generare esternalità negative come i costi di congestione. Inoltre bisogna ricordare che la struttura di costo nella pubblicità online si basa spesso su una strategia *pay-per-click*, che non costituisce un incentivo per gli inserzionisti a scegliere come *target* solo le grandi piattaforme. Per riassumere, gli effetti della rete sulla concorrenza sono contrastanti e sembrano generare solo un vantaggio temporaneo, mentre l'evidenza suggerisce che l'innovazione continua e la qualità del prodotto sono le chiavi per avere successo (Lerner 2014).

⁴ fenomeno per cui in statistica un modello si adatta a un campione perché ha un numero eccessivo di dati che permette quindi anche a modelli sbagliati di adattarsi ai dati disponibili.

3.5 I BIG DATA COME FONTE DI “PRICE DISCRIMINATION”

È la “Pratica di imporre prezzi diversi a consumatori diversi per l’acquisto dello stesso bene o di forme leggermente differenti” (Treccani). Se ci riferiamo agli sconti basati sulle diverse quantità acquistate allora siamo in presenza di discriminazione di secondo grado come ad esempio paghi uno prendi due; mentre la discriminazione di terzo grado consiste nel praticare un prezzo diverso a ogni gruppo di clienti diverso come ad esempio negli sconti per *over 65*. Le forme di discriminazione di secondo e terzo grado dei prezzi sono largamente utilizzate ed utilizzate nella società; mentre la discriminazione di primo grado, nota anche come discriminazione perfetta, è quella in cui viene fissato un prezzo diverso per ogni individuo, consentendo teoricamente alle imprese di appropriarsi di tutto il surplus del consumatore. In realtà è sempre stato considerato abbastanza utopico come processo, poiché risultava quasi impossibile conoscere il prezzo di riserva di ciascun consumatore per ogni bene.

Gli studiosi oggi concordano nel dire che l'analisi dei *big data* ha favorito il passaggio dalla discriminazione di secondo e terzo grado a quello di primo grado (Nuccio e Guerzoni 2018). Grazie ad un utilizzo massivo di dati di tipo comportamentale e multidimensionale che consentono la segmentazione dei consumatori, le piattaforme digitali possono avvicinarsi quanto più alla “*willingness to pay*”⁵ di gruppi molto piccoli di consumatori o addirittura dei singoli individui. Utilizzando i dati sugli abbonati Netflix, è stato dimostrato che l'aumento dei profitti reso possibile dai prezzi personalizzati è molto più elevato quando si utilizzano dati raccolti sui comportamenti di navigazione sul Web piuttosto che dati con informazioni demografiche (Shiller 2014).

Townley, Morrison e Yeung concludono che "gli effetti sul benessere dei consumatori sono ambigui" (2017). Quando i mercati sono competitivi, la discriminazione può generare elevati vantaggi, offrendo ai clienti con minore disponibilità prezzi più bassi e aumentando i prezzi per coloro che sono disposti a pagare di più. Al contrario, altri potrebbero sostenere che le informazioni sul comportamento dei consumatori danneggino gli stessi che potrebbero finire per pagare più di quanto farebbero in assenza della

⁵ è il prezzo massimo al quale o sotto di cui il consumatore sicuramente comprerà un’unità di prodotto

condivisione dei dati. L'argomento più forte di questi oppositori è che le imprese in questo modo si appropriano della maggior parte o di tutto il *surplus* del consumatore avvicinandosi il più possibile al prezzo di riserva. La contro argomentazione sottolinea che, in un mercato monopolistico con costi fissi particolarmente elevati, è solo grazie alla discriminazione dei prezzi che il mercato diviene redditizio e accessibile anche ad utenti con una bassa propensione all'acquisto (Armstrong, 2008). In altre parole, l'extra profitto consente alle imprese di servire tutti i segmenti, compresi i mercati di scarso interesse. Più in generale, citando quella che è la teoria classica, gli economisti ritengono che il benessere sociale, definito come la combinazione del surplus dei consumatori e quella dei produttori, potrebbe aumentare quando la discriminazione espande effettivamente la produzione totale (Varian, 1985) e ai consumatori con una disponibilità a pagare inferiore rispetto a quella che porta all'equilibrio dei prezzi venga offerto un prezzo inferiore. La conclusione secondo *Kerber* sarebbe che in presenza di consumatori sofisticati e lungimiranti maggiori informazioni aumenterebbero la concorrenza e l'innovazione, mentre mantenere le informazioni private potrebbe non migliorare il benessere sociale e abbassare i prezzi (2016).

Dobbiamo concentrarci sia sulla "*price dispersion*"⁶ che sulla riduzione dei prezzi. Sia la dispersione dei prezzi nei mercati digitali che la riduzione complessiva dei prezzi dovuta al commercio elettronico è stata documentata da Gorodnichenko e Talavera (2017) e da Willis (2004). La combinazione di queste due prove suggerisce che le imprese che operano nei mercati digitali perseguono la *price discrimination* abbassando i prezzi per i consumatori con una bassa disponibilità a pagare, e questo indica un probabile effetto positivo della discriminazione dei prezzi. Pertanto, possiamo concludere che la discriminazione dei prezzi ha un effetto positivo sui consumatori. Bisogna però ricordare che mentre l'effetto sui consumatori è positivo, la riduzione dei margini di profitto dovuta alla riduzione dei prezzi va a scapito dei profitti dei venditori. In effetti, la posizione dominante di una *digital platform* le consente di imporre al venditore una strategia di prezzo, come dimostra il caso dello scontro tra Amazon e gli editori tradizionali.

⁶ La dispersione dei prezzi si verifica quando venditori diversi offrono prezzi diversi per lo stesso bene in un determinato mercato

3.6 L'EFFETTO DEI BIG DATA SULL' INNOVAZIONE

Se in un quadro economico statico lo sfruttamento del potere di mercato porta alla perdita di benessere a spese dei consumatori, ciò non avviene se prendiamo come riferimento i guadagni in termini di efficienza dovuti alla presenza di innovazioni derivanti da extra profitti. In effetti, si potrebbe sostenere che gli utili in eccesso generati dal potere di mercato sono una condizione necessaria per affrontare gli ingenti investimenti in R&S richiesti dall'innovazione (Acs and Audretsch, 1987). L'approccio dei cosiddetti "regimi tecnologici" considera la tecnologia di base di un settore come il principale fattore che determina sia la struttura del mercato sia della performance innovativa. Distinguiamo quindi quei settori dove i livelli di innovazione tecnologica sono bassi, il grado di appropriabilità è elevato e l'accumulo di conoscenza è importante, detti "Schumpeter Mark II", qui tendiamo ad osservare alcuni grandi e stabili innovatori, dal caso opposto, ossia quello dei mercati dove possiamo osservare il cambiamento turbolento dei leader di mercato cosiddetti "Schumpeter Mark I". Malerba e Orsenigo (1997) hanno dimostrato che la maggior parte delle industrie può essere classificata in uno di questi due casi.

Negli anni '90, al tempo della bolla delle dot-com, l'economia di Internet era chiaramente un Schumpeter Mark I, in cui l'innovazione era introdotta da un afflusso costante di nuove società. Tuttavia, sembra che ora per lo sfruttamento delle opportunità tecnologiche siano necessari ingenti investimenti in ricerca e sviluppo, come dimostrato dalla recente intensificazione della ricerca di *Facebook*, *Alphabet* e *Amazon*. Nel 2013 *Alphabet* e *Amazon* sono entrati nella top 20 dei maggiori investitori in ricerca e sviluppo, raggiungendo rapidamente la prima e la seconda posizione nel 2017⁷.

⁷ Statista (2017) <https://www.statista.com/statistics/265645/ranking-of-the-20-companies-with-the-highest-spending-on-research-and-development/>.

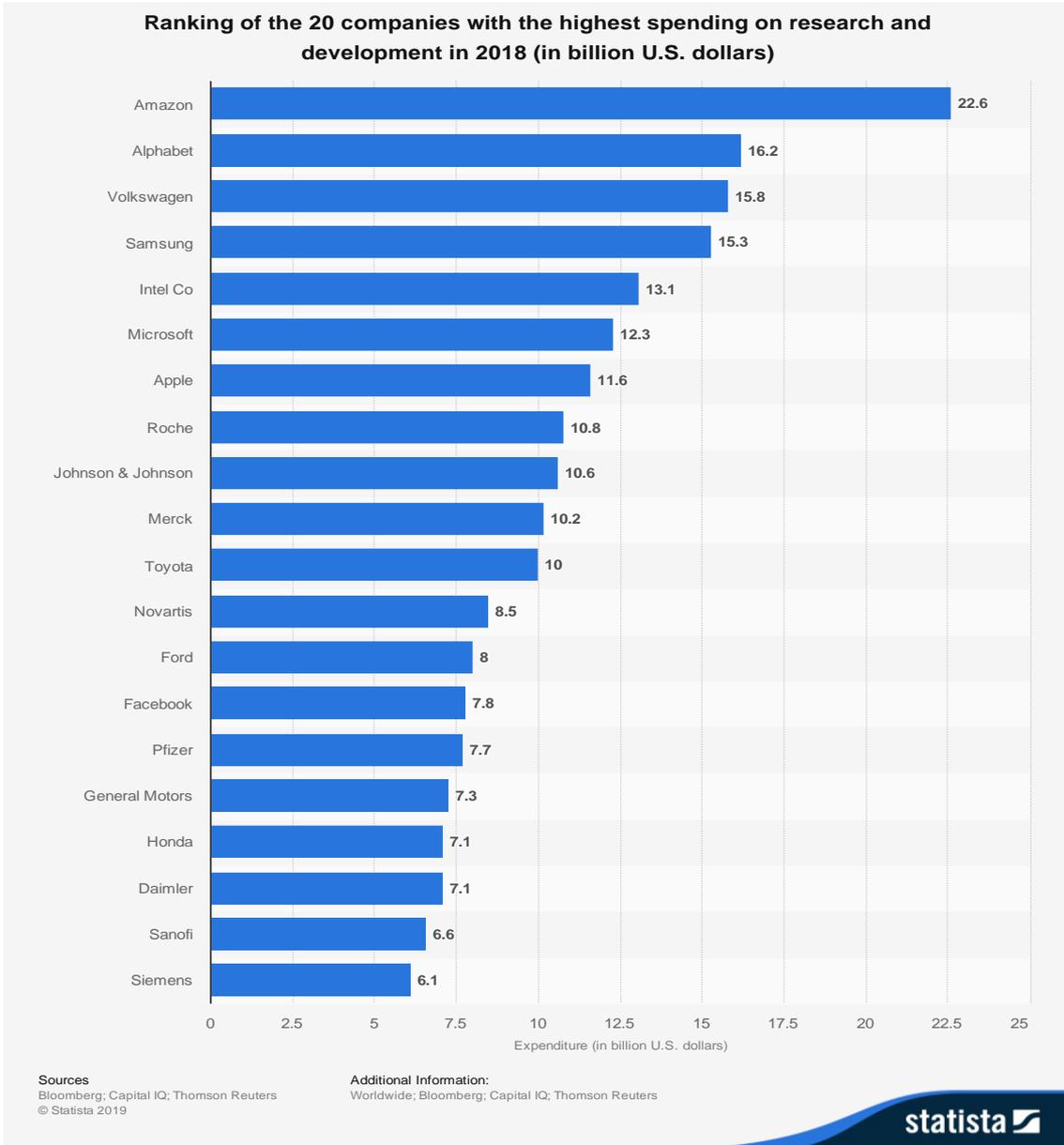


Figura 68

⁸ Spendere soldi in ricerca e sviluppo è il modo in cui un'azienda innova le proprie tecnologie e conduce ricerche per sviluppare nuovi prodotti e servizi. La ricerca e sviluppo consente alle aziende come Amazon di stare al passo con la concorrenza e lavorare per il futuro. La maggior parte della spesa in R&S nel mondo è in informatica ed elettronica, e gli Stati Uniti sono il paese leader per le spese in R&S nel mondo. Attualmente molte aziende stanno investendo nell'innovazione dell'intelligenza artificiale (AI), anche per ciò che riguarda le implicazioni etiche. Le principali aziende, come Amazon e Alphabet, si stanno gettando in maniera molto decisa verso questa direzione, e conducono ricerche sulle possibili applicazioni future. Con le aziende leader a livello mondiale che spendono miliardi di dollari in ricerca e sviluppo, il futuro dell'“AI” sicuramente progredirà rapidamente.

Inoltre, le conoscenze coinvolte nell'analisi dei big data e nel modello di business sono piuttosto stabili, rendendo il processo di accumulo delle conoscenze piuttosto cumulativo. Possiamo quindi affermare che la concentrazione del mercato nella “internet economy” è il risultato della tecnologia sottostante, che ha però permesso un altissimo ritmo nel cambiamento tecnologico che altrimenti sarebbe stato impossibile. Questa evidenza rende piuttosto difficili le decisioni delle autorità antitrust, in quanto devono monitorare e bilanciare attentamente quello che è l'impatto negativo dello sfruttamento dell'enorme potere di mercato che però ha come diretta conseguenza un'elevata efficienza dinamica e innovativa.

3.7 ABUSI E POSIZIONI DOMINANTI

Valutiamo qui il rischio di abuso di posizione dominante esaminando tre aspetti economici. In primo luogo, sosteniamo che le imprese non possono facilmente abusare della loro posizione dominante poiché nei mercati digitali vi è un alto grado di contendibilità. La natura stessa dei dati e le tecnologie in rapida evoluzione non garantiscono barriere all'ingresso a lungo termine.

In secondo luogo, il ruolo di coordinatori degli *incumbents* e la grande quantità di informazioni in loro possesso li mettono nella condizione di poter attuare politiche di discriminazione con i consumatori. Tuttavia è stato precedentemente dimostrato che i consumatori tendono a beneficiare delle pratiche discriminanti nei mercati online, poiché le aziende attuano tali pratiche abbassando i prezzi per i mercati delle fasce più basse, espandendo così la produzione, lasciando inalterato il benessere per i vecchi consumatori. In terzo luogo, evidenziamo che le leggi delle autorità *antitrust* tollerano lo sfruttamento del potere di mercato ma quando mira a migliorare il benessere per i consumatori, ad esempio investendo nel miglioramento dei prodotti esistenti o sviluppandone di nuovi. Nonostante le grandi aziende tecnologiche generino miliardi di dollari di entrate nette, raramente esse distribuiscono dividendi; a volte utilizzano profitti per riacquisti di azioni, ma principalmente utilizzano i fondi per investimenti in ricerca e sviluppo e per l'acquisizione di società high-tech (Molla 2017).

In questo contesto, è proprio la natura del progresso tecnologico a generare il potere di mercato e che solo con sforzi straordinari in ricerca e sviluppo con cui gli *incumbents*

possono fronteggiare i propri concorrenti e potenziali entranti. Inoltre, a causa dalle tecnologie emergenti che possono sconvolgere il mercato sono necessari ingenti e continui investimenti in ricerca e sviluppo. È rispettato quindi il principio della libera concorrenza in quanto le imprese se vogliono accaparrarsi posizioni dominanti lo possono fare solo nel momento in cui offrono un valore aggiunto per il consumatore.

I politici e gli studiosi che tendono a sottolineare il rischio del controllo del mercato da parte delle società di big data sono soliti sopravvalutare l'aspetto tecnologico rispetto al fattore umano. Il valore per le aziende dei big data si basa sulla capacità di estrarre conoscenza e non sulla sola semplice acquisizione e archiviazione. Ciò che non è abbastanza sottolineato nel paradigma dei big data è che "i dati, di per sé, hanno spesso un valore basso" (Rubinfeld e Gal, 2017); ciò che è richiesto non sono solo capacità tecnologiche o analitiche, ma competenze eterogenee e, soprattutto, un nuovo approccio all'innovazione continua. Molte aziende generano già molte informazioni sui propri clienti senza la necessità di big data e potrebbero persino essere in grado di elaborarle con tecniche avanzate *di data mining e machine learning*. Lo sforzo maggiore per diventare un'azienda basata sui dati è il raggiungimento di cambiamenti culturali e organizzativi, concentrandosi sul valore economico delle informazioni.

CAPITOLO QUARTO

POLITICAL MICROTARGETING: IL CASO CAMBRIDGE ANALYTICA

4.1 LA PROFILAZIONE DELL'UTENTE

Quando si parla di profilazione si fa riferimento a quell'insieme di attività realizzate attraverso processi, *software* ed algoritmi che raccolgono ed elaborano dati sugli utenti che utilizzano un determinato servizio per poi suddividerli in gruppi a seconda del loro comportamento.

In ambito commerciale è largamente impiegata per fornire servizi personalizzati come ad esempio la pubblicità comportamentale, è inoltre utile per ottenere analisi accurate dei potenziali clienti. Secondo il GDPR dell'Unione Europea, di recente entrato in vigore, per profilazione si intende “qualsiasi forma di trattamento automatizzato di dati personali consistente nell'utilizzo di questi ultimi per valutare determinati aspetti personali relativi a una persona fisica (...)”. Nella realtà, ai fini puramente commerciali si tende a fare uso di qualsiasi tipologia di informazione rinvenibile online creando una profilazione particolarmente dettagliata quasi sempre lesiva o al limite della normativa concernente la *privacy*. (Saetta, 2018)

È su questo tessuto che hanno trovato applicazione le tecniche di profilazione utilizzate dalla società di consulenza britannica Cambridge Analytica, che attraverso il *data mining* e l'analisi dei dati attuava comunicazioni strategiche per scopi elettorali. (Calia, 2018)

Tutto ebbe inizio nel 2008 quando Michal Kocinski, un'esperto di psicologia comportamentale, venne ammesso all'Università di Cambridge per lavorare ad un Dottorato di Ricerca nel Centro di psicomètria¹ del Campus. Lì, con l'aiuto del suo collega David Stillwell realizzò l'applicazione “*my personality*” che attraverso la

¹ branca della psicologia che cerca di misurare costrutti astratti e intangibili. Per svolgere questa operazione, un tratto come la personalità, viene scomposto in indicatori visibili e misurabili.

piattaforma Facebook proponeva agli utenti di compilare un questionario psicometrico basato sul modello “*Big five*”.

Tale modello impiega sofisticate metodologie statistiche per valutare i tratti della personalità in una modalità particolarmente accurata. Esso nella pratica si serve di cinque dimensioni delineate dal c.d. modello *OCEAN* (*openness, consciuosness, extroversion, agreeableness, neuroticism*) e che costituiscono la configurazione della personalità di ciascun individuo. Come hanno dimostrato già numerose ricerche tali metodi permettono di effettuare efficaci analisi predittive sul comportamento e sulle preferenze di ciascun utente facilitando, come nel caso della Cambridge Analytica, sia il marketing politico che quello elettorale.

Un'altra modalità per analizzare le caratteristiche psicologiche degli utenti è l'analisi psicografica che, sebbene sia molto efficace, avrebbe comportato dei costi estremamente ingenti se somministrato all'intero elettorato americano ed è proprio relativamente a tale problematica che Kosinski ha rivestito un ruolo chiave.

I risultati del lavoro “*Private traits and attributes are predictable from digital records of human behavior*” realizzato da quest'ultimo e dal suo collega Stillwell furono così sorprendenti da farlo considerare secondo la rivista “*Altmetrics*” come il quarto lavoro più influente di sempre. I due dottorandi sono stati capaci di dimostrare come analizzando alcuni specifici “*like*” particolarmente rappresentativi su *Facebook* si possa giungere ad un'analisi predittiva con cui definire attributi sui soggetti come ad esempio la religione o l'orientamento sessuale, ma anche informazioni sull'affiliazione ad un partito piuttosto che ad un altro.

Successivamente il *team* di Kosinski ha intrapreso delle nuove ricerche basate sulla possibilità di attribuire potere predittivo a qualsiasi altra traccia digitale come ad esempio l'attività fisica svolta da un individuo e registrata dal proprio *smartphone* oppure dalla musica ascoltata sul *device*.

Nel 2014, attraverso l'intermediazione del collega Aleksandr Kogan, la società “*Strategic communication laboratories*”, operante nel settore di marketing e comunicazione politica basata su modelli psicografici, acquisisce la metodologia *OCEAN* che verrà a sua volta utilizzata dalla società collegata Cambridge Analytica specializzata nel marketing elettorale.

Alexander Nix, amministratore delegato della Cambridge Analytica, ha più volte illustrato le modalità con cui la società riusciva, partendo da semplici tracce digitali, a condurre delle analisi predittive sulle preferenze degli elettori ed a fornire a questi ultimi dei feedback personalizzati volti ad influenzarne le scelte politiche. Tutto ciò è stato possibile grazie alla tecnologia “*big five*” che, sebbene abbia suscitato forti dubbi sulla sua correttezza morale, ha permesso la raccolta di informazioni personali di molteplici elettori americani attraverso il metodo della segmentazione psicografica. (Amenduni 2017)

4.2 POLITICAL MICROTARGETING

Attualmente i *social media* giocano un ruolo di fondamentale importanza durante le campagne elettorali. Ormai le comunicazioni attraverso questi ultimi possono essere personalizzate sulla base di molteplici informazioni che pervengono dagli utenti in rete e fornite anche dagli stessi volontariamente. Ciò ha comportato non solo un risparmio dei costi delle campagne elettorali ma soprattutto ha permesso ai soggetti politici di raccogliere una maggiore quantità di voti che altrimenti non avrebbero ottenuto.

Nonostante la portata innovativa di tale contesto, il marketing politico *online* è un fenomeno ancora poco conosciuto.

Il *microtargeting* di per sé consiste nel ricercare e combinare i diversi dati ed informazioni personali fornite dagli utenti al fine di ricavarne le preferenze politiche e successivamente nel sottoporre agli stessi dei messaggi personalizzati, il tutto grazie a piattaforme online. Attualmente ciò che manca è un vero e proprio studio sull’atteggiamento psicologico degli utenti con riguardo alle così dette “elezioni digitali”: spesso molti soggetti non sono a conoscenza del fatto di essere stati individuati come possibili *target* e se ne fossero stati a conoscenza probabilmente non avrebbero mai accettato di essere esposti a determinate informazioni.

I ricercatori Sanne Kruikemeier , Minem Sezgin e Sophie C.Boerman hanno condotto uno studio per comprendere quali fossero le risposte degli utenti ai messaggi politici personalizzati basati sulle informazioni da loro stessi fornite.

Gli studi condotti hanno dimostrato come gli annunci online possano avere effetti di gran lunga differenti a seconda degli utenti: Alcuni di loro li percepiscono come utili in quanto filtrano la grande quantità di informazioni presenti in rete permettendo di ottimizzare ed

agevolare i processi decisionali; Altri invece hanno espresso la propria preoccupazione soprattutto per quel che concerne la propria *privacy*.

Tuttavia i risultati di tale ricerca sono ancora molto relativi a causa di una mancanza di uno studio scientifico certo sulla persuasività degli annunci personalizzati.

Infatti il concetto di persuasione acquisisce importanza fondamentale quando si cerca di comprendere quali siano gli effetti e le conseguenze di tali messaggi: se gli utenti fossero pienamente consapevoli dei suggerimenti politici celati dietro agli annunci personalizzati probabilmente assumerebbero una condotta dissidente nei confronti di questi ultimi.

Ad esempio alcuni post di *Facebook* messi in rete da partiti politici appaiono nella nostra “*timeline*” semplicemente perché un nostro amico lo ha condiviso o commentato ma ciò non implica necessariamente che vi sia una connessione tra noi e quel determinato partito. Tali post sono molto simili a quelli normali eccetto che per l’etichetta “sponsorizzato” utile ad indicare il proposito commerciale di quel determinato messaggio. Ciononostante le ricerche condotte hanno dimostrato che tali etichette non sono sufficienti a rendere l’utente pienamente consapevole degli scopi degli annunci, sia perché spesso passano inosservate, sia perché gli stessi individui non sono a conoscenza del significato di tali avvertimenti.

Per tali ragioni si è condotto uno studio per cercare di comprendere se uno “*short training*”, con il proposito di informare gli utenti sulla pratica del microtargeting, potesse essere utile affinché questi ultimi possano essere pienamente consapevoli degli scopi politici dei messaggi personalizzati che ricevono.

Il risultato ha dimostrato che il più delle volte la consapevolezza degli scopi persuasivi celati dietro un messaggio ha comportato un’alterazione nella condotta degli utenti sia nei confronti del mittente che del messaggio stesso causato da una diversa capacità di elaborazione degli individui.

L’*Elaboration Likelihood Model* suggerisce due diverse tipologie di elaborazione:

- Attraverso il sistema centrale che implica una comprensione attenta e ponderata dello scopo delle informazioni
- Attraverso il sistema periferico che da all’individuo dei semplici indizi

Nell’esempio di Facebook sopra proposto se un individuo leggendo l’etichetta “sponsorizzato” fosse incentivato a controllare maggiormente la notizia elaborandola con

il sistema centrale, elaborerà conseguentemente il messaggio ed il suo contenuto in modo più attento e critico.

Nel caso in cui non venisse fornita alcuna informazione riguardante la natura commerciale o politica della notizia ci sarebbero invece più probabilità di elaborare il messaggio attraverso il sistema periferico del nostro cervello con una maggiore probabilità che l'utente non sia pienamente consapevole degli scopi persuasivi del messaggio stesso.

Tutto ciò implica che se un individuo comprendesse pienamente la capacità persuasiva del messaggio, essendogli noti gli scopi, considererà quel mittente non più affidabile.

In altre parole, se un messaggio politico personalizzato contenesse un preavviso sulla sua natura persuasiva indurrebbe gli utenti ad una elaborazione più critica dello stesso e ad una conseguente riduzione della fiducia riposta nel mittente ovvero nel partito politico in questione. (Liebert , 2015)

4.3 MICROTARGETING POLITICO ATTUATO DA CAMBRIDGE ANALYTICA

La Cambridge Analytica ha utilizzato una tipologia di *microtargeting* politico che implica l'utilizzo di tecnologie particolarmente avanzate come ad esempio il *microtargeting* psicografico.

Quest'ultimo consiste nel misurare la personalità degli elettori sulla base delle loro tracce digitali e successivamente di influenzarne le scelte attraverso l'invio di messaggi personalizzati.

Nella pratica Cambridge Analytica si è servita di annunci pubblicitari creati sulla base delle personalità dei diversi utenti e messi in rete attraverso la piattaforma Facebook, veicolo da quest'ultima prescelto al fine di raggiungere gli scopi di propaganda politica.

Nel 2017 Michael Kosinski , esperto in psicologia comportamentale presso l'università di Cambridge, ha condotto una serie di ricerche volte a dimostrare l'efficacia del *targeting* politico come strumento di propaganda digitale arrivando alla conclusione che si tratta di un vero e proprio mezzo di persuasione digitale di massa.

Lo studio effettuato da quest'ultimo attraverso delle campagne pubblicitarie diffuse su *Facebook* ha messo in luce l'impossibilità di proibire il *targeting* psicologico con scopi di propaganda politica proprio a causa del fatto che sono gli stessi utenti a fornire volontariamente informazioni personali anche se il più delle volte inconsapevolmente.

Nello specifico, per quel che concerne gli strumenti utilizzati da Cambridge Analytica per la raccolta dati ed al fine di influenzare le condotte degli elettori è possibile catalogarne almeno tre.

- *Cookies*: file inviati sul nostro computer al fine di rendere tracciabili le nostre ricerche in rete.

Si tratta di strumenti sempre più diffusi negli anni grazie all'impiego di tecniche capaci di trasformarli in veri e propri strumenti di marketing e di profilazione a fini pubblicitari.

- *Facebook Audience Insight*: funzionalità che permette di analizzare un pubblico selezionato manualmente o definito attraverso determinate caratteristiche personali. In particolare si fa riferimento a parametri come l'età, il sesso, la religione, la provenienza, il titolo di studio oppure a delle semplici attività svolte sulla piattaforma come ad esempio condivisione di determinati *post* e *like*. L'analisi dei dati risultante da tali suddivisioni viene utilizzata per inviare dei messaggi personalizzati ai gruppi di utenti selezionati ed è utilizzata sia per valutare l'ingresso in un determinato mercato che per valutare gli effetti di una campagna pubblicitaria già in atto.
- *Looklike Audiences su Facebook*: si tratta di uno strumento che ha permesso alla società Cambridge Analytica di trovare utenti che hanno caratteristiche simili ad utenti già profilati e di modulare i messaggi di propaganda politica anche in sulla base della personalità di questi ultimi. (Sisti, 2017)

4.4 IL CASO CAMBRIDGE ANALYTICA

La società di consulenza britannica Cambridge Analytica venne creata nel 2013 dal miliardario repubblicano Robert Mercer inizialmente per scopi che ruotavano attorno alle elezioni politiche statunitensi dello stesso anno che vedevano coinvolto il senatore repubblicano Ted Cruz.

L'attività della società si incentrava sostanzialmente nella ricerca di informazioni e dati sui consumatori ed utenti al fine di creare pubblicità mirate ed altri servizi in ambito sia politico che aziendale.

Dopo la campagna elettorale del senatore Cruz, la società ha intrapreso nuove collaborazioni politiche, acquisendo un ruolo di spicco sia nella campagna vincente sulla Brexit sia in quella del noto imprenditore Donald Trump.

Da quanto emerge dall'inchiesta svolta dal New York Times e dal The Guardian, Cambridge Analytica avrebbe raccolto milioni di profili Facebook degli elettori statunitensi per creare un potente *software* per prevedere ed influenzare la scelta alle urne, realizzando così una delle più grandi violazioni di dati mai realizzate nel mondo digitale. Al tempo dei fatti, l'informatore Christopher Wylie, ex dipendente di Cambridge Analytica rivelò al *The Observer* come la società britannica utilizzasse le informazioni personali ottenute senza un esplicito consenso degli individui per costruire un sistema in grado di profilare i singoli elettori per indirizzarli nelle scelte attraverso pubblicità politiche personalizzate.

Nel 2015 la conferma di quanto accaduto arrivò proprio grazie ad una dichiarazione rilasciata da Facebook che iniziò ad adottare delle misure volte a limitare i danni ed a proteggere le informazioni personali di oltre 50 milioni di elettori.

La raccolta di dati avvenne grazie alla creazione di un'applicazione chiamata "*thisisyourdigitallife*" ideata dal ricercatore dell'Università di Cambridge Aleksander Kogan. Egli attraverso una società da lui creata in collaborazione con la Cambridge Analytica, la Global Science Research, pagò milioni di utenti affinché questi ultimi si sottoponessero ad un test sulla personalità utilizzando come espediente una ricerca con finalità accademiche.

L'app creata da Kogan inoltre era in grado di raccogliere non solo i dati degli utenti sottoposti ad i quiz ma riusciva anche ad averne l'accesso e dunque a raccogliere le informazioni personali degli amici di Facebook di questi ultimi.

Una volta emerso il misfatto Facebook e Cambridge Analytica si sono ritrovati al centro di un'indagine sull'utilizzo illegale dei dati con finalità di marketing politico, svolta a livello Europeo da parte del British Information Commissioner's Office. Nello specifico l'oggetto dell'indagine verteva sull'utilizzo dell'analisi dei dati per scopi politici ed era stata avviata per comprendere come i partiti politici, le società di analisi di dati e le

piattaforme di *social media* utilizzassero ed analizzassero le informazioni personali degli elettori per realizzare il così detto *microtargeting* psicografico.

Pochi giorni dopo le dichiarazioni rilasciate da Facebook al The Observer, la stessa ha riservato la propria posizione legale affermando la sospensione dalla piattaforma delle inserzioni create da Cambridge Analytica in attesa di informazioni più chiare relative all'utilizzo improprio dei dati.

La vicenda ha suscitato un'indignazione tale da indurre il senatore democratico Mark Warner a proporre un "*Honest Ads Act*" al fine di regolare la pubblicità politica online allo stesso modo della televisione, della stampa e della radio tale da migliorare i controlli sulla raccolta di dati su vasta scala per il *targeting* politico.

Nonostante ciò l'amministratore delegato di Cambridge Analytica e la piattaforma Facebook hanno continuato a mantenere una linea difensiva affermando che si trattasse semplicemente di *fake news* e che nella sostanza la società non possedeva né utilizzava dati privati di Facebook illecitamente acquisiti.

Tuttavia le dichiarazioni rilasciate da questi ultimi furono immediatamente contraddette dall'ex dipendente della società Kogan il quale, per rendere più chiaro il quadro della vicenda, fornì un ampio *dossier* di prove al The Observer che a sua volta le sottopose all'esame dell'unità per la criminalità informatica della National Crime Agency ed all'Ufficio del Commissionario per le informazioni.

Tale *dossier* includeva prove come *e-mail*, fatture, contratti e trasferimenti bancari che rivelavano come oltre 50 milioni di profili, perlopiù appartenenti ad elettori statunitensi, erano stati prelevati dalla piattaforma Facebook realizzando una delle più grandi violazioni di dati dell'universo digitale.

Infatti, al tempo della violazione dei dati, l'informatore era dipendente della società Cambridge Analytica ma Facebook continuava ad affermare che egli ormai ricopriva un ruolo collaboratore di una società dallo stesso fondata già dal 2014.

Tuttavia anche tali affermazioni furono smentite: il dossier di prove fornito conteneva infatti una lettera inviata dai legali di Facebook allo stesso informatore Wylie nella quale gli era stato richiesto di distruggere tutti i dati in suo possesso acquisiti grazie alla Global Science Research, società fondata da Kogan per profilare tutti gli utenti.

Successivamente fu lo stesso Kogan ad affermare che la società da lui creata avrebbe ottenuto l'accesso ai dati degli utenti in modo legittimo e per le vie appropriate ma che in

seguito Facebook e Cambridge Analytica non hanno rispettato le regole sulla *privacy* fornendo tali informazioni anche a terze parti.

Ad ulteriore conferma di quanto accaduto, tra le prove consegnate al The Observer, è inoltre emerso un contratto risalente al 4 giugno 2014 che conferma che la Global Science Research, consociata di Cambridge Analytica, aveva siglato un accordo commerciale interamente basato sulla raccolta di dati dalla piattaforma Facebook.

Da quanto emerge, la Cambridge Analytica ha speso quasi 1 milione di dollari per raccogliere dati e profilare più di 50 milioni di utenti da abbinare alle diverse liste elettorali. Nella sostanza la società ha dunque utilizzato i risultati dei test sulla personalità ed i dati di Facebook per creare un algoritmo capace di analizzare i singoli profili Facebook e determinare i tratti salienti della personalità legati alle preferenze in ambito politico.

Al tempo dei fatti, 50 milioni di profili rappresentavano circa un terzo degli utenti di Facebook nordamericani attivi e quasi un quarto dei potenziali elettori statunitensi.

Nonostante le evidenze, la società Cambridge Analytica ha continuato ad affermare che il contratto stipulato con la Global Science Research prevedeva che Kogan, prima di entrare in possesso dei dati personali degli utenti, avrebbe dovuto ottenere il consenso informato per la raccolta dei dati, per cui non vi era alcun motivo di credere che quest'ultimo non lo avrebbe fatto.

Inoltre un'affiliata della società, la SCL Elections, secondo quanto riportato da Cambridge Analytica, avrebbe lavorato in quel periodo in collaborazione con Facebook per assicurarsi che non fossero state intenzionalmente violate le condizioni sulla *privacy* fornendo oltretutto una dichiarazione firmata attestante che tutti i dati ed i derivati per cui mancava un consenso esplicito erano stati istantaneamente eliminati.

Nonostante le innumerevoli evidenze probatorie, la vicenda si è conclusa con una dichiarazione di Cambridge Analytica con la quale afferma che in ogni caso nessuno dei dati in suo possesso è stato utilizzato al fine di realizzare un *microtargeting* politico. (Cadwalladr and Graham-Harrison, 2018)

4.5 SVILUPPI DELLA VICENDA E REGOLAMENTAZIONE

Il 2 Maggio 2018 la società Cambridge Analytica ha dichiarato ufficialmente la cessazione dell'attività e l'inizio delle procedure di insolvenza in Gran Bretagna.

Tutti i dipendenti hanno in seguito dovuto ottemperare all'ordine di restituire tutti i materiali e dispositivi di cui erano in possesso per permettere agli inquirenti di svolgere ulteriori indagini.

Ciononostante, nell'annunciare la cessazione dell'attività, Cambridge Analytica ha continuato a difendersi dalle accuse sostenendo che si trattasse di pratiche politiche e commerciali ampiamente accettate come "componente standard della pubblicità.

Dopo lo scandalo generato dalla violazione dei dati la piattaforma Facebook ha attuato delle misure volte a monitorare in modo più efficiente le applicazioni di terze parti all'interno del *social media*. Tali misure hanno richiesto sia la sospensione di più di 200 applicazioni create con lo scopo di raccogliere e registrare i dati degli utenti iscritti alla piattaforma sia ulteriori controlli condotti dai responsabili della sicurezza e della *privacy* volti ad assicurarsi che nessuna informazione venisse prelevata illegittimamente.

Anche Google, in seguito a Facebook, ha dichiarato di aver adottato un sistema di verifica volto ad analizzare gli annunci degli inserzionisti elettorali negli Stati Uniti al fine di prevenire condotte illegittime.

Nello specifico chiunque voglia acquistare degli spazi pubblicitari a fini politici o commerciali è tenuto a fornire un documento d'identità valido volto ad attestare la cittadinanza o residenza legittima nel Paese, così come richiesto dalla legge.

Nel 2016 si è poi finalmente assistito ad una delle più grandi rivoluzioni in tema di protezione dei dati grazie all'introduzione del Regolamento UE 2016/679, noto come *General Data Protection Regulation* (GDPR). Il testo di legge è stato approvato dal Parlamento Europeo il 14 Aprile 2016, ed entrato ufficialmente in vigore in tutti gli stati membri il 25 Maggio 2018.

Il Regolamento è stato introdotto per tutte quelle aziende che hanno come attività principale quella di raccolta ed elaborazione dei dati personali dei cittadini europei sia online che offline, sia fuori che all'interno della comunità Europea.

L'articolo 4.1 stabilisce l'oggetto del regolamento definendo come "dato personale": qualsiasi informazione riguardante una persona fisica identificata o identificabile («interessato»); si considera identificabile la persona fisica che può essere identificata, direttamente o indirettamente, con particolare riferimento a un identificativo come il nome, un numero di identificazione, dati relativi all'ubicazione, un identificativo online o a uno o più elementi caratteristici della sua identità fisica, fisiologica, genetica, psichica, economica, culturale o sociale."

Tuttavia si ritiene che la definizione debba essere interpretata in un senso più ampio ed includere dunque non solo i classici dati sensibile come sesso, indirizzo o numero telefonico ma anche identificativi online come ad esempio i cookies, la mail o la geo-localizzazione.

Il Regolamento inoltre definisce come "consenso dell'interessato": "qualsiasi manifestazione di volontà libera, specifica, informata e inequivocabile dell'interessato, con la quale lo stesso manifesta il proprio assenso, mediante dichiarazione o azione positiva inequivocabile, che i dati personali che lo riguardano siano oggetto di trattamento"

È necessario che vi sia un consenso formalizzato per ogni operazione con il conseguente diritto di ciascun utente di poter prendere visione dei dati personali raccolti che lo riguardano e di poter revocare in qualsiasi momento il consenso precedentemente dato.

Ulteriore novità apportata nel medesimo ambito riguarda l'autorizzazione che il soggetto titolare dei dati deve fornire affinché questi ultimi possano essere utilizzati.

In alcune circostanze poi il Codice della *Privacy* permette il silenzio assenso ovvero dei casi in cui è sufficiente che non vi sia un'opposizione esplicita affinché determinati dati possano essere trattati.

Come emerge dagli art 4.7 e 4.8 nell'ambito del trattamento dei dati si dovranno definire due figure principali:

- Il "titolare del trattamento": "la persona fisica o giuridica, l'autorità pubblica, il servizio o altro organismo che, singolarmente o insieme ad altri, determina le finalità e i mezzi del trattamento di dati personali; quando le finalità e i mezzi di tale trattamento sono determinati dal diritto dell'Unione o degli Stati membri, il

titolare del trattamento o i criteri specifici applicabili alla sua designazione possono essere stabiliti dal diritto dell'Unione o degli Stati membri”

- Il “responsabile del trattamento”: “persona fisica o giuridica, l’autorità pubblica il servizio o altro organismo che tratta dati personali per conto del titolare del trattamento”.

L’obiettivo del GDPR è quello di introdurre un meccanismo di responsabilizzazione dei soggetti implicati per far sì che il Regolamento sia correttamente ed efficacemente applicato.

Per consentire un trattamento dei dati lecito e trasparente viene, in primo luogo, richiesto che i titolari ed i responsabili dei dati tengano due registri:

- Il registro del titolare del trattamento, che deve contenere: anagrafica del titolare stesso, di un contitolare se presente, del rappresentante e del titolare alla protezione dati; le finalità del trattamento; le categorie degli interessati a cui fa capo il dato; eventuali termini per la cancellazione automatica del dato; un’eventuale descrizione generale delle misure di sicurezza tecnico-organizzative
- Il registro del titolare del trattamento, in cui sono presenti: l’anagrafica dei responsabili del trattamento; la descrizione delle categorie dei trattamenti effettuati; opzionalmente la descrizione delle misure di sicurezza intraprese

In secondo luogo, è necessario che si adottino delle misure di sicurezza consistenti nel notificare le violazioni o il trattamento illegittimo dei dati entro 72 ore dalla scoperta all’autorità di controllo, salvo che queste ultime non comportino rischi.

“La notifica dovrebbe includere:

- Una descrizione della natura della violazione dei dati personali compresa, ove possibile, la categoria ed il numero approssimativo degli interessati in questione, nonché le categorie ed il numero approssimativo di registrazioni dei dati personali;
- La comunicazione con il nome ed i dati di contatto del responsabile della protezione dei dati;
- Una descrizione delle probabili conseguenze della violazione dei dati personali;

- Una spiegazione delle misure adottate (o in via di adozione) da parte del titolare del trattamento per porre rimedio alla violazione o per attenuarne i possibili effetti negativi.”

Il nuovo GDPR richiede ai titolari del trattamento non solo di garantire il rispetto delle norme ma anche di dimostrare che queste ultime vengano attuate efficientemente. Di conseguenza è possibile introdurre il concetto di *accountability*, ovvero una vera e propria responsabilità sull’attuazione delle misure richieste dal Regolamento che comporta una valutazione di volta in volta dei possibili rischi e dei conseguenti effetti delle violazioni verificatesi da parte del titolare.

Altra figura di rilievo in tale ambito è quella del responsabile della protezione dei dati che all’interno del GDPR prende il nome di *Data Protection Officer*. Tale figura è incaricata di accertarsi che il titolare dei dati o il responsabile al trattamento dei dati osservino gli obblighi imposti dal Regolamento.

Un altro aspetto fondamentale del nuovo testo è l’introduzione del “diritto all’oblio”. In altre parole il titolare dei dati personali trattati ha il diritto di poter richiedere al titolare del trattamento la cancellazione delle informazioni personali in possesso di quest’ultimo nei seguenti casi:

- Quando non siano più necessari rispetto alle finalità per le quali sono stati raccolti o altrimenti trattati;
- Quando l’interessato revoca il consenso su cui si basa il trattamento e se non sussiste altro fondamento giuridico per il trattamento;
- Quando l’interessato si oppone al trattamento e non sussiste altro motivo legittimo prevalente per procedere al trattamento
- Quando i dati devono essere cancellati per adempiere ad un obbligo legale previsto dal diritto dell’Unione Europea o dello Stato membro cui è soggetto il titolare del trattamento;
- Quando i dati personali raccolti appartengono ad un minore e dunque ad un soggetto non pienamente consapevole dei rischi derivanti dal trattamento.

Per quanto riguarda la “portabilità dei dati” è richiesto che tutti i dati raccolti e posseduti dai social network siano facilmente accessibili permettendo anche la possibilità di modificarli o cancellarli.

Novità sono state introdotte anche nell'ambito della modalità di registrazione e fruizione di taluni servizi online comportando pertanto una variazione relativa alla "profilazione dell'utente" definita di seguito dal Regolamento come:

"qualsiasi forma di trattamento automatizzato di dati personali consistente nell'utilizzo di tali dati personali per valutare determinati aspetti personali relativi a una persona fisica, in particolare per analizzare o prevedere aspetti riguardanti l'andamento professionale, la situazione economica, la salute, le preferenze personali, gli interessi, l'affidabilità, il comportamento, l'ubicazione o gli spostamenti di detta persona fisica. Nela caso di marketing diretto, l'interessato avrà sempre diritto di opporsi alle attività di profilazione"

Un caso a sé stante riguarda i dati provenienti da soggetti di età inferiore a 16 (o, in casi specifici, di 13) che, essendo meno consapevoli dei rischi, delle conseguenze e delle misure di salvaguardia nonché dei loro diritti in relazione al trattamento dei dati personali, necessitano di una specifica protezione.

La condizione affinché i dati in questione possano essere trattati si traduce in un'autorizzazione da parte del genitore salvo che non si tratti di servizi forniti direttamente al minore. L'autorizzazione è necessaria soprattutto se l'intento è quello di procedere alla profilazione dell'utente per finalità di marketing o politiche. (Hombberger , 2018)

4.6 SANZIONI APPLICABILI PER LA VIOLAZIONE DEL REGOLAMENTO

- Sanzioni fino a 10 milioni di euro o 2% del fatturato totale annuo di un'impresa per: violazione delle condizioni applicabili al consenso dei minori, trattamento illecito di dati personali che non richiede l'identificazione dell'interessato, mancata o errata comunicazione di un *data breach* all'autorità nazionale competente, violazione dell'obbligo di nomina del DPO (Data Protection Officer), mancata applicazione di misure di sicurezza;
- Sanzioni fino a 20 milioni di euro o 4% del fatturato totale annuo di un'impresa per: inosservanza di un ordine imposto dall'autorità nazionale competente e trasferimento illecito di dati personali ad un destinatario in un Paese terzo.

4.7 PANORAMA ATTUALE

Nonostante lo scandalo provocato Cambridge Analytica e dalle conseguenti misure di protezione attuate in ambito di *Data Protection*, ci sono ancora molte riserve concernenti le modalità di operare di molte piattaforme online che spesso si trovano al confine legittimo e consentito per legge.

Facebook, Google ed altri sono in continua ricerca di nuovi modi per fare pubblicità soprattutto attraverso l'utilizzo dell'intelligenza artificiale, realtà virtuali ed il neuro marketing. A tal proposito si è infatti soliti parlare di *Internet of Things* ovvero una costante raccolta di dati e monitoraggio delle nostre attività che consentono ai servizi 4 di offrire delle prestazioni sempre più mirate ed efficienti.

Il potere e la precisione dei big data tende spesso ad essere una problematica sottovalutata pur essendo spesso correlata ad innumerevoli scandali nel mondo digitale.

Infatti il Sole 24ore, in un recente articolo datato 19 agosto 2019, pone nuovamente Facebook sotto i riflettori denunciandone le pratiche scorrette.

Secondo l'autorevole giornale, la piattaforma avrebbe pagato centinaia di collaboratori esterni affinché questi ultimi trascrivessero le note audio tra gli utenti Messenger in modo da ricavarne un testo scritto e quindi dati strutturati facilmente analizzabili.

Dopo la rivelazione dell'agenzia Bloomberg, i soggetti implicati sono stati costretti ad ammettere la pratica ed a porvi fine. Sebbene la *policy* sull'uso dei dati personali non faccia alcun cenno alle note audio ed alla possibilità che questi ultimi possano essere ascoltati da terze parti, alla vicenda non ha fatto seguito alcuna ipotesi di reato in quanto Facebook avrebbe fatto trascrivere solo le chat vocali degli utenti che hanno acconsentito a tale opzione.

Le finalità sono ancora discusse ma tra le possibilità si crede vi sia quella di testare quanto l'intelligenza artificiale utilizzata per trascrivere l'audio fosse performante.

CONCLUSIONE

Arriva a questo punto è bene soffermarsi e trarre quelle che sono le conclusioni a cui si è giunti con il presente elaborato.

Ho analizzato la letteratura in materia di big data con l'intento di dare rilevanza a questo fenomeno per le imprese, contestualizzandolo nell'era della digitalizzazione.

Per fare ciò è stato necessario inizialmente inquadrare il fenomeno big data all'interno del mondo delle imprese. Ciò è stato fatto nella prima parte dell'estratto con lo scopo di far emergere la *liaison* che si è consolidata dagli anni '70 ad oggi proprio tra le imprese ed i dati. Emblematico è l'utilizzo della catena di valore per spiegare quanto le imprese apprezzino effettivamente tale risorsa, preservandola ed analizzandola con tecniche che richiedono investimenti ragguardevoli e sforzi notevoli.

Addentrando maggiormente nella letteratura il lavoro mostra quello che è il passaggio successivo alla raccolta, allo *storage* ed all'analisi dei dati, ossia la possibilità effettiva di creare valore da questo processo. Il secondo capitolo offre dunque una dissertazione sul tema del marketing e come esso sia cambiato con l'avvento dei big data. Spunti di riflessione sono dati dalla professionalizzazione del *programmatic advertising* che grazie all'utilizzo di flussi continui di informazioni oggi ha la possibilità di creare profili dei propri utenti molto specifici e arrivare agli stessi con una proposta di valore altamente specifica. Il *programmatic advertising* come si evince nei passaggi successivi ha fatto molta strada evolvendosi a tal punto di consentire alle imprese di fornire il giusto tipo di *impression* al consumatore a seconda di dove esso si trovi nel percorso di acquisto.

Un secondo aspetto che hanno rivoluzionato i big data con il proprio avvento sono le cosiddette *digital platforms*, la forte concentrazione dei mercati nelle mani di pochi colossi del *tech* a cui abbiamo assistito, è stata resa possibile da questo nuovo modello di business ed al supporto continuo da parte dei dati alle piattaforme stesse. A tale proposito l'elaborato si sofferma sulla *sharing economy* spiegando questo modello di economia alternativo a quelli tradizionali, che nasce proprio dalla possibilità di sfruttare i Big data. Al termine di questo capitolo si è voluto mostrare con analisi sulle barriere, sulle discriminazioni di prezzo, sull'innovazione, e sugli abusi di posizioni dominanti cosa sia cambiato per quanto riguarda l'aspetto della concorrenza tra le imprese.

Il caso con cui si va a chiudere questo lavoro riguarda la società di consulenza britannica Cambridge Analytica, che ha utilizzato queste tecniche per fini commerciali, non sempre

in maniera limpida e a tutela degli utenti, creando un precedente che ha dato il via alla proliferazione di una robusta legislazione in materia di privacy del consumatore.

In conclusione le ricerche svolte mostrano chiaramente che i dati rappresentano per il mondo delle imprese un fattore chiave per assicurarsi non solo una fetta importante del mercato per alcune, ma perfino la sopravvivenza per quelle che hanno voluto e vogliono continuare a competere nonostante le sfide della digitalizzazione.

BIBLIOGRAFIA

Capt 1

Bardowicks, B., & Busch, O. (2013). Diskussionspapier: Programmatic Advertising.
bvdw.org

Bernard Marr, 2015, Why only one of the 5 Vs of big data really matters

Boccellato Piero, 2018, Information security, la spesa globale nel 2018 sarà di 93 miliardi di dollari

GSMA, 2018, the data value chain

Han Hu, Yonggang Wen, Tat-Seng Chua, Xuelong Li, 2014 Toward Scalable Systems for Big Data Analytics: A Technology Tutorial

Mandelli A. 2017, big data marketing creare valore nella platform economy con dati, intelligenza artificiale e IoT, p. 67.

Wedel Michel & Kannan P. K., 2016, Marketing Analytics for Data-Rich Environments

Capt 2

Heimann, T. (2013). Supply-Side-Plattformen

Lamb Patrick Dawson Michael, 2015, Enhanced Success with Programmatic Social Advertising

Wedel Michel & P.K. Kannan, 2016, Marketing Analytics for Data-Rich Environments

PageFair, 2016, mobile ad blocking report, pagefair.com

Smith, MD, Telang, R (2016) Streaming Sharing Stealing.

Wedel Michel & P.K. Kannan, 2016, Marketing Analytics for Data-Rich Environments

Capt 3

Armstrong, M (2008) Price discrimination.

Barbieri Elisa, di Tommaso Marco Rodolfo, 2012 - Dizionario di Economia e Finanza
Treccani

Familiar, B, Barnes, J (2017) Business in real-time. In: Business in Real-Time Using
Azure IoT and Cortana Intelligence Suite.

Gertz Oliver e McGlashan Deirdre, 2016, Consumer-Centric Programmatic Advertising

Haucap, J, Heimeshoff, U (2014) Google, Facebook, Amazon, eBay: Is the Internet
driving competition or market monopolization?

Kerber W (2016) Digital markets, data, and privacy: competition law, consumer law and
data protection.

Lerner, AV (2014) The role of 'big data' in online platform competition

Molla R, (2017), "Tech companies spend more on R&D than any other companies in the
U.S." Recode

Nuccio Massimiliano, Guerzoni Marco, 2018, Big data: Hell or heaven? Digital platforms
and market power in the data-driven economy

Rubinfeld, DL, Gal, MS (2017) Access barriers to Big Data.

Shiller, BR (2014) First degree price discrimination using big data

Simon, H (1971) Designing organizations for an information-rich world. In: Greenberger, M (ed) Computers, Communications, and the Public Interest

Townley, C, Morrison, E, Yeung, K (2017) Big data and personalized price discrimination in EU competition law.

Van Alstyne, 2016, how platform business are transforming strategy.

Varian, HR (1985) Price discrimination and social welfare.

Capt 4

Homberger Luca, 2018, il GDPR in 10 passi:il nuovo regolamento Europeo sulla privacy, CWS digital solutions

Liebert Mary Ann, 2015, Cyberpsychology , behavior , and social networking

Sisti Alessandro, 2017, Digital transformation war: Retailer tradizionali VS Giganti dell'e-commerce

SITOGRAFIA

Capt 1

eBay study, 2013, how to build trust and improve the shopping experience

Villanova university, What is big data?, 2019
<https://www.villanovau.com/resources/bi/what-is-big-data/>

Capt 3

Ayers Ryan, 2017, 5 Ways Big Data is Fueling the Sharing Economy,
<https://www.smartdatacollective.com/5-ways-big-data-is-fueling-the-sharing-economy/>

Global big data conference, 2016, The Sharing Economy - What It Is, Examples, And How Big Data, Platforms And Algorithms Fuel,
<http://www.globalbigdataconference.com/news/130013/the-sharing-economy---what-it-is-examples-and-how-big-data-platforms-and-algorithms-fuel.html>

<https://www.glossariomarketing.it/significato/ad-exchange>

Larrarin Daniele, 2018, Sharing Economy, il dilemma continua: rivoluzione vera o "economia dei lavoretti"? <https://www.digital4.biz/executive/sharing-economy-il-dilemma-continua-rivoluzione-vera-o-economia-della-miseria/>

Washington state university, How The Sharing Economy Is Transforming Business,
<https://onlinemba.wsu.edu/blog/how-the-sharing-economy-is-transforming-business/>

Capt 4

<https://www.cnbc.com/2018/03/20/cambridge-analytica-claimed-firm-ran-digital-campaign-for-trump-report-says.html?&qsearchterm=data%20mining%20cambridge%20analytica>

<https://www.valigiablu.it/cambridge-analytica-big-data-trump-facebook/>

<https://www.cnn.com>

<https://brunosaetta.it/privacy/>

<https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election>