UNIVERSITA' DEGLI STUDI DI PADOVA

DIPARTIMENTO DI SCIENZE ECONOMICHE ED AZIENDALI "M.FANNO"

CORSO DI LAUREA IN ECONOMIA

PROVA FINALE

"OPPORTUNITA' E RISCHI DEL DATA BROKERING"

RELATORE:

CH.MO PROF. FABIO MANENTI

LAUREANDO/A: CRISTINA FANTE

MATRICOLA N. 1172684

ANNO ACCADEMICO 2019 – 2020

La candidata, sottoponendo il presente lavoro, dichiara, sotto la propria personale responsabilità, che il lavoro è originale e che non è stato già sottoposto, in tutto o in parte, dalla candidata o da altri soggetti, in altre Università italiane o straniere ai fini del conseguimento di un titolo accademico. La candidata dichiara altresì che tutti i materiali utilizzati ai fini della predisposizione dell'elaborato sono stati opportunamente citati nel testo e riportati nella sezione finale 'Riferimenti bibliografici' e che le eventuali citazioni testuali sono individuabili attraverso l'esplicito richiamo al documento originale.

"You may not know them, but data brokers know you"

Edith Ramirez, 2014

INDICE

INTRODUZIONE

Il presente elaborato analizza la realtà dei *data brokers,* figure ancora poco conosciute che stanno avendo però un grande impatto nella nostra vita quotidiana. Ogni giorno infatti ci viene richiesta una miriade di informazioni, in particolare durante le nostre attività di navigazione *online.* Bisogna però prestare molta attenzione all'uso che si fa e si può fare di questi dati. Un uso corretto è "*un'azione indispensabile nei processi decisionali di imprese, istituzioni e, sempre più, anche dei singoli cittadini*" (AGCOM, 2018, p.2); un uso corretto delle informazioni, infatti, sta alla base della concorrenza e dello sviluppo delle aziende e può creare grande valore per l'economia mondiale, incrementando la quantità e la qualità dei prodotti e dei servizi offerti.

I dati non sono un concetto del tutto nuovo: già nel 2006 infatti, Clive Humby, l'artefice del successo della carta fedeltà (*clubcard*) della catena di supermercati britannica Tesco, sosteneva che "*i dati sono il nuovo petrolio*" (Casali A., 2019). Mikko Hypponen, Chief Research Officer di F-Secure disse però che, "*come il petrolio ha portato prosperità ma anche problemi, lo stesso faranno i dati*" (Sattler J., 2017); e per quanti vantaggi possano portare, infatti, bisogna imparare a gestirli per "estrarre" beneficio dalla loro raccolta. Per fare ciò ci sono i *brokers*.

Nel **primo capitolo** di questo elaborato verranno descritte le funzioni principali dei *data brokers*, le fonti di acquisizione dei dati, a chi vengono rivenduti e per quali scopi.
Successivamente, nel **secondo capitolo**, avvalendosi di alcune teorie, si cercherà di dimostrare se la presenza di maggiori informazioni sul mercato porti ad un aumento o ad una riduzione della competizione. E' un dibattito molto attuale e ancora aperto: ad oggi, non ci sono risposte chiare in merito.
Nel **terzo capitolo** si tratterà il tema della *privacy*, argomento di scottante attualità dal momento che le dimensioni del mercato comportano un enorme flusso di dati la cui origine sono gli individui, e questi dati contengono informazioni personali che dovrebbero essere protette con regolamenti stringenti e pene severe in caso di violazione degli stessi.
Il **quarto ed ultimo capitolo,** infine, presenterà il caso *Facebook-Cambridge Analytica*, che ha fatto scandalo in quanto informazioni personali degli utenti di *Facebook* sono state utilizzate con scopi ben diversi da quelli indicati all'atto di iscrizione al *social network*.

CAPITOLO 1

I DATA BROKERS[1]

Quotidianamente, le persone effettuano attività *online* che rivelano informazioni personali sul loro conto. Queste azioni comprendono ad esempio l'utilizzo dello *smartphone*, l'iscrizione ad un giornale, l'acquisto di un prodotto, la risposta a questionari per ricevere un *coupon* e, sempre più, l'utilizzo dei *social media*. Mentre i consumatori sono impegnati in queste attività, le entità con cui interagiscono collezionano informazioni su di loro e le forniscono o vendono ai *data brokers*.

In questo primo capitolo cercherò di approfondire la figura del *data brokers*, quali sono le loro principali attività, l'impatto del loro operato nella quotidianità e come si muovono nel mercato.

1.1 Chi sono i data brokers

I *data brokers*, anche definiti *information broker* o *information reseller*, sono aziende che collezionano informazioni *online* da fonti pubbliche, le aggregano, le interpretano e le analizzano per poi venderle sul mercato, costituendo parte integrante dell'economia dei *Big Data.* Le informazioni raccolte provengono da una vasta gamma di fonti con obiettivi molto diversi tra loro, tra cui verificare l'identità di un individuo, creare campagne di *marketing* personalizzate e rilevare frodi. Si tratta di dati personali che vengono successivamente organizzati in "profili", per essere venduti ad imprese che, ad esempio, intendono realizzare campagne pubblicitarie mirate o offrire servizi personalizzati. Poiché in genere ciascuna fonte fornisce pochi elementi sulle attività dei consumatori, i *data brokers*, oltre a riunire tutti questi elementi, li scambiano con altri *brokers* per formare profili più completi della vita del consumatore. L'attività di aggregazione e analisi dei dati può essere molto invasiva, consentendo di raggruppare gli individui in base all'etnia, al reddito, allo stile di vita, agli *hobby* e persino alle condizioni di salute e dedurne quindi gli interessi. Si evince che i consumatori, pur fornendo consapevolmente i propri dati alle varie fonti pubbliche, sono ignari dell'esistenza di tali organizzazioni e soprattutto della varietà di pratiche in cui sono coinvolti.

Più avanti in questo elaborato (**Capitolo 3: Privacy e Legislazioni**), verrà approfondito il problema della trasparenza dell'operato dei *data brokers*. La *Federal Trade Commission* (da

[1] Se non specificato diversamente, le informazioni contenute in questo capitolo si riferiscono al documento "DATA BROKERS. A Call for Transparency and Accountability" (FTC, 2014). Disponibile su https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf.

qui, FTC), nel Dicembre del 2012, ha iniziato uno studio sulle pratiche dei *data brokers* distribuendo a nove *brokers* (Acxiom, Corelogic, Datalogix, eBureau, ID Analytics, Intelius, PeekYou, Rapleaf, and Recorded Future) direttive identiche per la compilazione di *report* specifici volti a raccogliere informazioni sulle loro pratiche di collezione e utilizzo dei dati e sugli strumenti forniti ai consumatori per controllare queste pratiche. **Le aziende coinvolte sono tutte ubicate nel territorio statunitense; i dati raccolti e le analisi effettuate si riferiscono quindi al territorio americano.** Tali direttive richiedevano informazioni dettagliate sulle azioni dei *data brokers*, con particolare riferimento alla natura e alle fonti delle informazioni collezionate, all'uso e alla diffusione delle stesse, ai consumatori coinvolti e all'eventuale possibilità di questi ultimi di accedere e correggere i propri dati ed eventualmente bloccarne anche la diffusione.

Secondo questo studio della FTC, i *brokers* dei dati collezionano informazioni da fonti commerciali, governative e altre pubblicamente disponibili. Nello sviluppare i loro prodotti, i *data brokers* usano sia i dati "grezzi" che ottengono dalle fonti, come ad esempio nome, indirizzo, età, sia dei dati dedotti da questi: per esempio, un *broker* dei dati può dedurre che un individuo con una patente nautica abbia un interesse per le barche (FTC, 2014, p. 19).
Dai dati raccolti, i *brokers* inferiscono gli interessi dei consumatori e possono quindi dividerli in categorie. Alcune categorie possono sembrare "innocue", come ad esempio "proprietario di un cane"; altre però, che si focalizzano primariamente su etnia e livello di reddito, oppure età e situazione sentimentale, o infine condizioni di salute, risultano categorie particolarmente sensibili.
Osservando a quali siti i consumatori si registrano, i *brokers* possono risalire a quali sono i loro interessi e *hobby* al di fuori del web e di conseguenza possono ipotizzare su quali siti i consumatori navigheranno.

1.2 Fonti: acquisizione dei dati

Sempre secondo lo studio della FTC, nessuno dei nove *brokers* coinvolti colleziona informazioni direttamente dai consumatori. Le principali fonti di acquisizione dei dati sono le seguenti: fonti governative, fonti commerciali e altre fonti pubblicamente disponibili.

1.2.1 <u>Fonti governative</u>:

a. La maggior parte dei nove data brokers osservati ottengono informazioni direttamente da fonti del governo. Per esempio, l'Ufficio del Censimento americano fornisce informazioni demografiche, come l'etnia, l'età, il livello di

istruzione, il reddito ecc. degli abitanti di una particolare area della città. Inoltre, fornisce informazioni geografiche come strade, indirizzi, confini, distretti scolastici e di voto. L'amministrazione della previdenza sociale fornisce informazioni quali il Death Master File, nel quale si possono trovare i nomi, i *Social Security Number* (SSN) e le date di morte dei consumatori. Inoltre, altre agenzie federali e internazionali forniscono informazioni relative a liste di controllo antiterrorismo.

b. Oltre ad utilizzare dati ricevuti direttamente dal governo federale, i governi statali e locali offrono molte informazioni, tra le quali licenze professionali e ricreative, informazioni sugli aventi diritto di voto, archivi giuridici, ecc. Bisogna però sottolineare che la maggior parte dei *brokers* non ottiene informazioni direttamente dai governi statali, bensì da altri *data brokers* che assumono persone per andare negli uffici locali e compilare le informazioni oppure che abbiano conoscenze negli uffici che diano loro il permesso di avere accesso automatico a queste informazioni.

1.2.2 Fonti commerciali

Quasi tutti i *data brokers* coinvolti acquistano informazioni da fonti commerciali ad ampio raggio. Ad esempio, i *brokers* ottengono dati dettagliati e specifici sugli acquisti da parte dei rivenditori e delle aziende di cataloghi. Queste informazioni possono riguardare il prodotto/servizio acquistato, l'ammontare speso per l'acquisto, la data dell'acquisto e il tipo di pagamento utilizzato. Alcuni *brokers* ottengono informazioni dagli editori sulle tipologie di abbonamenti venduti; informazioni come nome, indirizzo e indirizzo mail da siti che richiedono l'iscrizione per erogare servizi; informazioni sulle transazioni da aziende di servizi finanziari, ecc. La maggior parte dei dati proviene da altri *data brokers* che ottengono informazioni da compagnie telefoniche, da venditori di auto o anche da questionari, come sondaggi di *marketing* o *contest* che i clienti compilano online o offline.

1.2.3 Fonti pubblicamente disponibili

Più della metà dei *brokers* hanno raccolto dati pubblicamente disponibili, come ad esempio numero di telefono, data di nascita e altre informazioni che gli individui postano in Internet, nei loro *blog* o nei loro *account social*. Alcuni di loro acquisiscono informazioni dai *social media*, come ad esempio *LinkedIn*,

piattaforma in cui le persone tendono a non limitare l'accesso ai propri dati personali (FTC, 2014, p.13); *LinkedIn* diventa un'importante fonte, senza restrizioni.

1.3 Principali tipologie di dati

I nove *brokers* interpellati dalla FCT sono interessati a dati relativi in particolare ai settori del *marketing* (prodotti di *marketing*), della mitigazione del rischio (prodotti per la mitigazione del rischio) e della ricerca di persone. Di queste tre categorie quelli maggiormente richiesti sono senza dubbio i prodotti di *marketing*, seguiti dalla mitigazione del rischio.

1.3.1 Prodotti di marketing

Per entrare più nel dettaglio, i prodotti di marketing sono i dati che le aziende richiedono per creare: a) campagne di ***marketing* diretto**, che comprendono la posta, il *telemarketing* e il *marketing* per e-mail; b) servizi di ***marketing online***, che comprendono il *marketing* su Internet, sui cellulari o attraverso la televisione; c) ***marketing* analitico**. Queste tre categorie danno la possibilità ai clienti dei *data brokers* di creare messaggi di *marketing* personalizzati.

a ***Marketing* diretto.** In base alle informazioni ricevute, lo studio della FTC ha identificato due categorie di prodotti di marketing diretto: "*data append*" e liste di *marketing*.

I prodotti "*data append*" (aggiunta di dati) aiutano le aziende ad apprendere di più sui loro consumatori; le aziende richiedono ai *data brokers* informazioni aggiuntive (per esempio numero di telefono o indirizzo e-mail) rispetto alle informazioni di base (nome e indirizzo del consumatore), da utilizzare nelle campagne di *telemarketing* o *marketing* per e-mail.

Le liste di *marketing* identificano i consumatori che condividono certe caratteristiche, ad esempio tutte le persone con almeno due figli, famiglie che hanno un fumatore tra i componenti, ecc. (FTC, 2014, p.25). Il cliente del *broker* individua le caratteristiche che vorrebbe trovare nei suoi consumatori e il *data broker* gli fornisce una lista di individui con quelle caratteristiche. Le liste possono limitarsi ai nomi e ai numeri di telefono per le campagne di *telemarketing*, oppure alle e-mail per le campagne via e-mail; se i clienti

richiedono dati più robusti per adattare le proprie campagne di *marketing*, i *brokers* possono includere nelle liste altri elementi descritti nei prodotti "*data append*".

b ***Marketing online.*** Il *marketing online* è stato suddiviso in tre categorie: *targeting* per registrazione, *targeting* per collaborazione e *onboarding*.

Targeting per registrazione. I *brokers* possono aiutare i siti che richiedono registrazione a promuovere i prodotti in maniera più efficace attraverso esperienze personalizzate sul consumatore. Ad esempio, se un sito di viaggi vuole proporre un particolare prodotto ai propri utilizzatori può inviare al *broker* una lista di utenti registrati e il *broker* fornisce al sito *web* una lista con gli interessi di viaggio di quegli specifici utilizzatori (FTC, 2014, p.26).

Targeting per collaborazione. Se nel *targeting* per registrazione i clienti dei *brokers* sono i siti dove i consumatori si iscrivono, nel *targeting* per collaborazione ci sono due tipologie di clienti: il sito che richiede registrazione e un inserzionista che sta cercando un sito dove inserire pubblicità personalizzata. Il sito e l'inserzionista danno al *broker* rispettivamente una lista degli utilizzatori e dei clienti. Nessuna delle due parti può avere accesso alle informazioni personali dei clienti dell'altra parte, solo il *broker* può avere accesso a tutte le informazioni e a questo punto può analizzarle per decidere se l'inserzionista può fare pubblicità sul sito o meno.

Onboarding. L'*onboarding* si riferisce ad un processo con il quale un *data broker* inserisce dati ricavati offline in un *cookie* per permettere agli inserzionisti di "colpire" consumatori praticamente ovunque sul *web*. L'*onboarding* prevede tre passaggi (Figura 1.1):

- Segmentazione
- Abbinamento
- Raggiungimento del consumatore

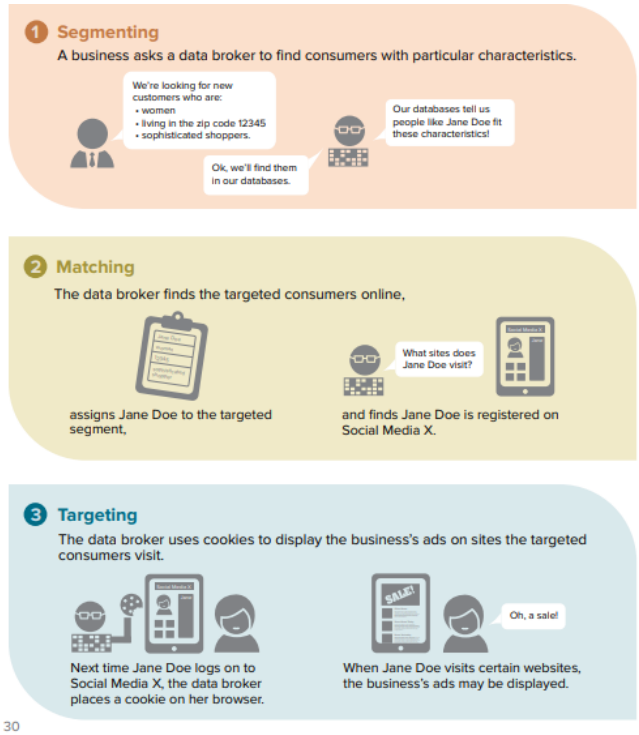

Fig. 1.1: Onboarding. Fonte: Federal Trade Commission, 05/2014, p.30

Il processo di *onboarding* inizia quando un cliente chiede ad un *data broker* di trovare dei consumatori con determinate caratteristiche (*segmentazione*). Il broker può aiutare il cliente sia comparando la propria lista di potenziali consumatori con quella del cliente, sia dando a quest'ultimo libero accesso alla propria lista per trovare nuovi consumatori.

Il secondo passaggio è l'*abbinamento*, durante il quale i *brokers* trovano i consumatori identificati durante il processo di segmentazione *online*. Per trovare i consumatori *online*, i *brokers* stipulano dei contratti con i siti per comprare le liste di utenti registrati. Queste liste poi vengono comparate con i consumatori identificati nel processo di segmentazione per trovare degli abbinamenti; una volta trovati, si allegano a quel consumatore tutti gli elementi o i segmenti associati con quel tipo di consumatore.

L'ultimo passaggio è la concentrazione sui clienti che sono stati trovati *online* (*raggiungimento del consumatore*). Per fare ciò, il *broker* deve inserire un *cookie* sul *browser* del consumatore quando il sito notifica che il consumatore è entrato nel sito; il *cookie* include le informazioni che il *broker* ha allegato al profilo del consumatore ma non include altre informazioni come nome, indirizzo

o indirizzo e-mail. Una volta posto il *cookie*, la pubblicità resta presente sul browser del consumatore fino a quando il cookie resta sul browser.

c **Il *marketing* analitico** può essere un modo per prevedere i possibili comportamenti dei consumatori. Tra le altre cose, i prodotti analizzati offerti dai *brokers* danno la possibilità ai clienti di mirare in maniera più accurata le campagne pubblicitarie, raffinare i messaggi dei prodotti e delle campagne e guadagnare informazioni utili sulle abitudini e le preferenze dei consumatori. Come parte dell'analisi, i *data brokers* possono aiutare i propri clienti a modellare gli esiti attesi di strategie di *marketing*, permettendo ai clienti di pubblicizzare meglio i propri prodotti ai consumatori. Questi prodotti analizzati sono solitamente basati su algoritmi che considerano centinaia di migliaia di dati, tra i quali dati storici provenienti dai clienti e dati raccolti direttamente dal governo, da fonti pubbliche e da fonti commerciali.

1.3.2 Prodotti di mitigazione del rischio.

I prodotti di mitigazione del rischio si dividono in: verifica d'identità e individuazione di frodi.

Verifica d'identità. In generale, i prodotti di verifica dell'identità assistono i clienti nella conferma delle identità degli individui. Ad esempio, le banche usano questi prodotti per rispettare i requisiti di verifica dell'identità "Conosci il tuo cliente" oppure come aiuto per scoraggiare le frodi quando un consumatore inizia una transazione (FTC, 2014, p.32). Questi prodotti sono offerti in diversi formati. In primo luogo, i *data brokers* associano ai clienti delle banche un punteggio numerico che indica il livello di rischio associato alla transazione. Se il punteggio è elevato, il codice esplicativo può affermare che il *Social Security Number* (SSN) fornito dal consumatore è associato ad un individuo deceduto, o l'indirizzo è associato ad una frode o all'indirizzo di una prigione, che il SSN sia stato usato molto frequentemente in un breve periodo di tempo o che il SSN sia attribuito ad un indirizzo diverso da quello presentato dal consumatore. In secondo luogo, i brokers offrono ai clienti un prodotto a quiz che di solito include domande alle quali i consumatori dovrebbero essere in grado di rispondere facilmente come ad esempio "Che indirizzi e-mail hai utilizzato?" o "Quando è il compleanno di tua madre?". In terzo luogo, i brokers offrono un

format "*match/no match*", che fornisce una conferma che le informazioni fornite dal consumatore combacino con quelle nei file dei *brokers*.

Individuazione di frodi. Alcuni *brokers* vendono i loro prodotti per aiutare i clienti a identificare o ridurre le frodi. I prodotti per l'individuazione di frodi possono assistere i clienti nella verifica dell'affidabilità o della veridicità delle informazioni che un consumatore presenta. I prodotti dei *data brokers* possono anche assistere le aziende che hanno avuto violazioni di dati, analizzando l'andamento per determinare se c'è stato un uso improprio delle informazioni personali violate.

1.3.3 Dati per la ricerca di persone

In misura molto ridotta, sempre secondo lo studio della FTC, alcuni *brokers* forniscono informazioni per la ricerca di persone (*people search*) che può essere mirata al monitoraggio delle attività di dirigenti aziendali o alla ricerca di vecchi amici o vecchi documenti. Generalmente queste informazioni sono richieste da singoli individui e il prodotto può essere fornito dai *brokers* gratuitamente o a fronte di un modesto compenso.

1.4 Aziende clienti

La tabella (Fig.1.2) riporta l'elenco delle principali categorie di clienti dei nove *brokers* coinvolti nello studio della FTC e le principali categorie di prodotti richiesti. Come esplicitato al paragrafo 1.3 i prodotti maggiormente richiesti sono quelli di *marketing*.

	Direct Marketing	Online Marketing	Marketing Analytics	Identity Verification	Fraud Detection	People Search
Alternative Payment Providers[i]				x	x	
Attorneys & Investigators	x					
Automotive Industry	x	x	x			
Consumer Packaged Goods Manufacturers[ii]	x	x	x			
Data Brokers	x	x	x	x	x	
Educational Institutions	x			x	x	
Energy/Utilities	x					
Government Entities	x		x	x	x	x
Hospitality/Travel/ Entertainment	x	x	x			
Individual Consumers						x
Insurance Companies	x		x	x	x	
Lenders/ Financial Services Firms	x	x	x	x	x	x

	Direct Marketing	Online Marketing	Marketing Analytics	Identity Verification	Fraud Detection	People Search
Marketing/ Advertising Firms	x	x	x	x	x	x
Media	x		x			x
Non-profit Entities/ Political Campaigns	x	x		x	x	
Pharmaceutical Firms	x		x			x
Real Estate Services	x				x	x
Retail Companies	x	x	x	x	x	x
Technology Companies[iii]	x	x	x			x
Telecom Companies[iv]	x		x	x	x	

[i] Alternative Payment Providers include companies who provide consumers with alternative methods of payment rather than traditional methods such as checks or credit cards.

[ii] Consumer Packaged Goods Manufacturers include companies that manufacture items that consumers use and have to replace frequently, such as food and beverages, apparel, and household products.

[iii] Technology Companies include hardware companies, software companies, Internet companies, and other technology companies.

[iv] Telecom Companies include telephone, mobile, cable and satellite television providers, and other telecommunication companies.

Figura 1.2: Clienti per tipo di prodotto e settore. Fonte: Federal Trade Commission, 05/2014, p.39

Il rapporto tra *brokers* e clienti è subordinato ad un contratto in cui vengono chiaramente stabilite le condizioni di utilizzo dei dati. La stipula del contratto è solitamente preceduta da una pratica di *screening*, durante la quale il *broker* si informa sul potenziale cliente, facendo ricerche, osservandone le caratteristiche ed eventualmente incontrandolo, prima di decidere se vendergli i propri prodotti. Queste pratiche variano al variare del tipo di prodotto (e.g., *marketing* o mitigazione del rischio), del tipo di dati (e.g., informazioni sullo stile di vita) e del tipo di cliente (e.g., banche o *retailer*). Nello studio, ad esempio, un *broker* ha riportato che non vende informazioni a clienti che fanno parte di certe industrie, come la pornografia, l'investigazione privata, la vendita illegale di droghe o armi, ecc. (FTC, 2014, p. 41).
Oltre al processo di *screening*, i *brokers* possono sottoporre alle aziende un questionario che determina se il potenziale cliente è un'entità legittima e se utilizzerebbe i dati rispettando la legge.
I clienti possono essere sottoposti anche a verifica dei conti.

Infine, i *brokers* che offrono prodotti per la mitigazione del rischio e per il *marketing* firmano contratti che definiscono cosa è permesso e cosa è vietato fare con i dati scambiati. È proibito, ad esempio il riutilizzo o la rivendita dei dati senza permesso, la decodificazione degli stessi, l'utilizzo illecito o che viola le linee guida di settore.

CAPITOLO 2

I DATA BROKERS E LA COMPETITIVITÀ DEL MERCATO

Acquisire informazioni sui consumatori è ormai un'attività obbligatoria per la maggior parte delle aziende. Ma la disponibilità di maggiori informazioni aumenta o diminuisce la concorrenza?

2.1 Il caso ScanTrack

Un caso che ho trovato interessante è quello discusso da Heli Koski (2018) che riguarda le due più importanti aziende finlandesi nel settore del food. Queste due aziende hanno deciso volontariamente di non usufruire più dei servizi dei *data brokers*.

L'articolo suggerisce che, dopo la fine del rapporto con i *brokers*, le due aziende abbiano goduto di minor competizione e abbiano potuto applicare prezzi più alti ai propri prodotti, anche se nel lungo periodo non si sa se l'assenza di *brokers* abbia effettivamente giovato all'azienda (Koski H., 2018).

La sempre più ampia presenza e importanza dei *data brokers* e degli algoritmi che vengono utilizzati per monitorare il comportamento dei consumatori porta ad un aumento della trasparenza del mercato, e anche ad una maggior possibilità di scambio di informazioni tra aziende concorrenti.

Le due aziende in questione sono la Ruokakesko (da qui in poi Kesko) e la Suomen Osuuskauppojen Keskuskunta (da qui in poi SOK), due aziende che coprono rispettivamente il 34 e il 40% circa del commercio alimentare finlandese. Entrambe le aziende usavano ScanTrack, un servizio di AC Nielsen, il più importante *data broker* nell'industria alimentare, che opera in più di 100 paesi. Il sistema di raccolta dati di AC Nielsen si basava sulle informazioni fornite dai *bar code* delle casse dei negozi, sugli acquisti dei clienti. Nel febbraio 2007, la *Finnish Competition and Consumer Authority* (FCCA) iniziò, su "indicazione" di Kesko e SOK, un'indagine sui potenziali problemi di antitrust legati all'utilizzo di ScanTrack. Tra il 2007 e il 2008 infatti entrambe le aziende hanno abbandonato l'utilizzo di ScanTrack, "prevedendo" la decisione di FCCA che di fatto ha riconosciuto che Kesko e SOK, tramite il sistema ScanTrack, stavano violando il *Finnish Competition Act* e l'articolo 81 della Comunità Europea, in materia di leggi sulla concorrenza. La chiusura dei rapporti tra le due aziende e AC Nielsen ha comportato la fine di ScanTrack in Finlandia (Koski H., 2018).

2.2 Articoli correlati

Il fenomeno *data brokers* è un fenomeno piuttosto recente e non ci sono ancora studi conclusivi sugli impatti dei *brokers* sul mercato, ma solo diverse teorie.

Lo studio di Bounie, Dubus e Waelbroeck (si veda Koski H., 2018, p.4) è uno dei primi ad esplorare gli impatti competitivi di un *broker* che, per massimizzare i propri profitti, decide di vendere dati alle aziende che competono nel mercato produttivo. Il loro modello teorico suggerisce che quando nel mercato è presente un *broker* che vende informazioni a due aziende concorrenti, le aziende si troveranno in una situazione di dilemma del prigioniero, nel senso che entrambe le aziende acquisteranno informazioni pur sapendo che la concorrenza aumenterà e i profitti saranno inferiori. Questo stato è considerato l'equilibrio di Nash, in quanto le aziende otterranno profitti minori rispetto ad una situazione in cui nessuno acquista informazioni dai *brokers* (o non ci sono *brokers* sul mercato), ma, in questo modo, non rischiano di essere ulteriormente svantaggiate nel caso in cui solo l'altra azienda li acquistasse.

L'autrice continua dicendo che, se da un lato, possedere più informazioni dà alle aziende la possibilità di discriminare sul prezzo ed estrarre quindi più surplus dai consumatori, portandole a guadagnare profitti maggiori, dall'altro, un mercato più trasparente, più ricco quindi di informazioni, dà ai consumatori la possibilità di confrontare i prezzi e questo può facilitare la concorrenza, dato che le aziende sono incentivate a fissare prezzi minori. In mercati oligopolistici la discriminazione dei prezzi porta infatti ad una competizione più intensa, a profitti minori per le aziende e ad una crescita di benessere dei consumatori data dal calo dei prezzi. (Koski H., 2018). L'autrice sottolinea comunque che la letteratura economica non fornisce risposte chiare alla domanda se il possesso di maggiori informazioni sui consumatori aumenti o diminuisca la competizione tra le aziende. Anche gli studi di Bounie, Dubus e Waelbroeck del 2019 sono in linea con quanto appena riportato. Gli autori mostrano che i *brokers* non ritengono profittevole vendere tutte le informazioni ad imprese che si trovano in concorrenza. Vendere tutte le informazioni ad imprese che si trovano in concorrenza infatti aumenterebbe la competizione e successivamente diminuirebbe la loro possibilità a pagare per quei dati.

Il modello considera due dimensioni: la quota di mercato dei *brokers* e la porzione di mercato nella quale competono. Quando i *brokers* hanno quote di mercato simmetriche, collezionano meno informazioni che però vendono a più aziende rispetto ad una situazione di monopolio, con conseguente aumento di competizione tra le aziende. Quando le quote di mercato dei *data*

brokers sono asimmetriche, con l'aumentare della dimensione del mercato competitivo, uno dei due *brokers* colleziona e vende più informazioni rispetto ad una situazione di monopolio e la quota di clienti è maggiore rispetto ad un mercato monopolistico (Bounie D., Dubus A., Waelbroeck P., 2019).

Un altro impatto che viene citato nell'articolo di partenza è la possibilità di collusione. Un consistente scambio di informazioni, infatti, può facilitare la collusione tra aziende dato che le aiuta a coordinarsi per alzare i propri prezzi portando ad entrambe profitti più che competitivi (*supracompetitive prices*). Lo scambio di informazioni aiuta anche le aziende a controllare se i concorrenti rispettano i prezzi di collusione concordati (Koski H., 2018). Mehra (si veda Koski H., 2018, p. 11) ritiene che gli algoritmi aumentino la trasparenza e l'accuratezza con cui le aziende possono prevedere le reazioni dei concorrenti e rispondere velocemente con variazioni di prezzo. Ezrachi e Stucke (si veda Koski H., 2018, p. 11) continuano dicendo che questi algoritmi permettono alle aziende di prevedere le reazioni dei concorrenti e le loro strategie. Grazie a questi fattori i concorrenti possono far pagare prezzi sopra il livello concorrenziale (*supracompetitive)* e riescono a mantenere la collusione (Koski H., 2018).

Infine, l'articolo riassume gli effetti dello scambio di informazioni in un mercato oligopolistico in due effetti opposti: un aumento della concorrenza e una riduzione dei prezzi o un aumento della collusione tra imprese e l'aumento dei prezzi fino a farli diventare *supracompetitive* (Koski H., 2018).

L'articolo elenca infine gli effetti in tre ipotesi:
Ip.1: La concorrenza tende ad essere più debole e i prezzi dei prodotti più alti quando le aziende scambiano informazioni tramite un broker rispetto all'assenza di scambio (o assenza di broker sul mercato).
Ip.2: La concorrenza tende ad essere più forte e i prezzi dei prodotti più bassi quando le aziende acquistano informazioni da un broker rispetto all'assenza di acquisto (o assenza di broker sul mercato).
Ip.3: La concorrenza tende ad essere più forte e i prezzi dei prodotti più bassi quando solo una delle aziende concorrenti acquista informazioni da un broker rispetto all'assenza di broker. (Koski H., 2018).

CAPITOLO 3

PRIVACY E LEGISLAZIONI

Nei capitoli precedenti abbiamo avuto modo di conoscere che tutto ciò che facciamo online lascia una traccia e che ci sono agenti pronti a raccogliere tutte le informazioni che condividiamo. Su questo aspetto ho cercato di delineare qual è l'impatto dei vari attori coinvolti in questo processo e di come esso può essere migliorato, controllato, regolamentato. Le figure chiave sono chiaramente: i consumatori, i *brokers*, le aziende e gli organismi di controllo.

3.1 Utenti-consumatori

Il grado di consapevolezza degli utenti in merito alla cessione delle proprie informazioni è un aspetto di vitale importanza (AGCOM, 2018, p.58). Spesso, quando scarichiamo una APP o ci registriamo su un sito, accettiamo le condizioni di servizio senza prestare attenzione a cosa è scritto in questi documenti, a volte per non chiarezza degli stessi, a volte perché ormai viviamo in una società frenetica e non c'è più il tempo/l'interesse di leggere qualsiasi tipo di istruzioni/normative sulla *privacy*. *"Le persone spesso rivelano dettagli sulla loro vita personale senza rendersene conto"* (Deulkar D., Gupta P., 2018). Deulkar e Gupta (2018) danno quattro suggerimenti per cercare di ovviare a questo problema:

- Non rivelare completamente la propria identità. Dire tutta la verità sul proprio conto è obbligatorio solo se si compilano dei moduli del governo o se è richiesto dalla legge. In caso contrario, ad esempio durante l'iscrizione a Facebook, non è obbligatorio che le informazioni siano completamente corrette. Per evitare di arricchire i *database* dei *brokers* con informazioni non richieste è utile anche evitare di riempire gli spazi che non sono obbligatori.
- Essere a conoscenza del fatto che i siti usano i *cookies* per sapere dove navighiamo. Uno dei passaggi più importanti per cercare di proteggere la nostra *privacy* è l'eliminazione dei *cookies* di tanto in tanto. Cambiare le impostazioni sulla *privacy* infatti non fermerà la raccolta di dati ma eviterà che i *brokers* possano seguire le nostre tracce online.
- Fare il *log out* dai *social media* quando si naviga in rete. Un esempio che tutti conosciamo sono le pubblicità di prodotti o siti sui quali abbiamo precedentemente navigato, che compaiono all'improvviso quando stiamo navigando su *Facebook* o altri *social media*.

- Usare protezioni virus aggiornate. Avere sempre una protezione dai virus aggiornata può aiutare a prevenire l'ingresso di *malware*/virus che potrebbe danneggiare o rubare dati privati dal nostro sistema.

Oltre alla maggiore consapevolezza, i consumatori dovrebbero avere la possibilità di poter cancellare *(opt-out)* le proprie informazioni dai *database* delle aziende, e di conseguenza dei *brokers*, se non vogliono che queste siano usate con scopi diversi dalla semplice registrazione ad un sito, per esempio.

Una delle paure dei consumatori, come citato sopra, è quella di essere vittime di frodi; c'è infatti timore che alcune aziende possano usare le informazioni personali in modo sbagliato, per pianificare truffe nei confronti dei consumatori. Tra i rischi percepiti vi sono l'utilizzo da parte delle *internet company*, direttamente o tramite la cessione a terzi, delle proprie informazioni senza il consenso e la possibilità di subire furti di identità o di ricevere pubblicità indesiderata (Deulkar e Gupta, 2018) (AGCOM, 2018).

3.2 Brokers

In merito alla tutela dei dati e al rispetto della *privacy*, le aziende di *brokering* sono sempre più sensibili. Un esempio è l'azienda americana Acxiom, una realtà che si descrive come "un'azienda che rende possibile il miglioramento delle esperienze dei consumatori combinando dati, tecnologia, etica e idee per creare le basi per costruire un *business* basato sui consumatori e che offre un sistema basato sui dati che permette alle aziende di trattare le persone come persone, come loro si aspettano di essere trattate".
Il video[2] di presentazione dell'azienda riassume bene tutti questi punti soffermandosi sul fatto che *"**people are still people**"* e *"know your customer to deliver personalized, marketing experiences that matter, **through the ethical use of data**"*.
Acxiom è un esempio di come molte aziende di raccolta dati stiano focalizzando parte del proprio operato sul rispetto della *privacy*. Vedremo poi infatti che essere più trasparenti nei confronti dei clienti ripaga con clienti più sicuri e fedeli.

Un fattore da tenere presente è il rischio di attacchi informatici. Può infatti succedere che una cattiva gestione dei dati o una loro incontrollata diffusione non dipenda da scorretti

[2] Fonte: https://www.acxiom.com/

comportamenti dei *brokers* ma dal fatto che questi possano essere stati rubati, come successe, ad es. nel caso ChoicePoint, qui descritto.

3.2.1 Il caso ChoicePoint

L'industria dei *data brokers* non era ancora così conosciuta finché, nel 2005, molti casi di violazioni di dati *(data breaches)* furono portati alla luce. Il primo fu il caso di truffa di dati nei confronti di ChoicePoint, un'azienda di *brokering* americana. Uno studio di Otto P. N., Anton A. I. e Baumer D. L. (2007), riporta come la violazione di massa negli archivi di ChoicePoint venne considerata un punto nodale in quanto da quel momento l'industria di *brokering* e il tema della *privacy* e della sicurezza di informazioni personali hanno suscitato l'interesse collettivo e da parte del Congresso degli Stati Uniti.

Il *business* di ChoicePoint consisteva nella vendita di dati di credito a valutatori di rischio in ambito assicurativo. Otteneva informazioni personali sui consumatori, quali nome, date di nascita e storie di credito che rivendeva a più di 50,000 *business* (FTC, 2006). La Fig. 3.2 ci dà un'idea della provenienza e della destinazione dei dati che passavano per ChoicePoint e anche degli obiettivi della vendita di quei dati.

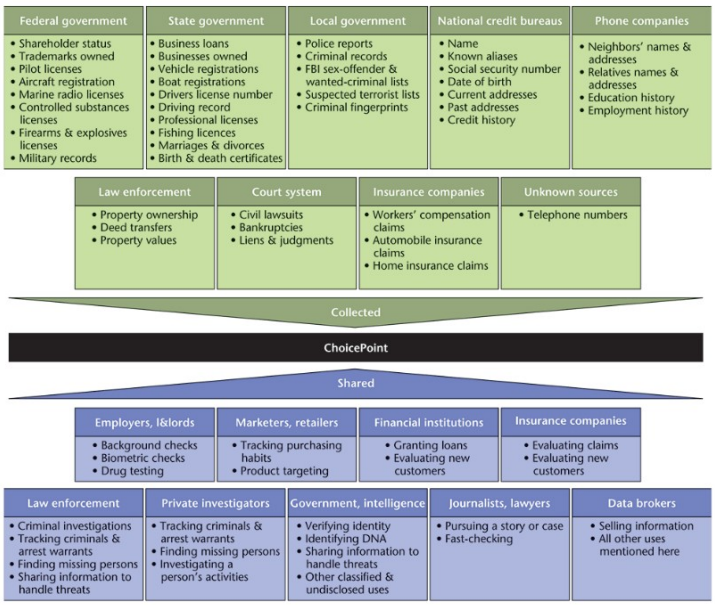

Fig. 3.1: The ChoicePoint business. Fonte: *"The ChoicePoint Dilemma: How Data Brokers Should Handle the Privacy of Personal Information",* 2007.

Il 14 Febbraio 2005 MSNBC.com, un canale televisivo statunitense, riportò la notizia che dei truffatori, che si erano finti *business* legittimi, erano entrati nei *database* di ChoicePoint e che potevano essere stati rubati i dati di 35,000 californiani. Fu chiaro fin da subito che non si parlava solo di abitanti della California e infatti alla fine del 2005 ChoicePoint dichiarò che le informazioni di circa 163,000 americani erano state violate. In realtà la truffa era iniziata già nel 2003 quando i ladri avevano acquistato licenze di *business* false, fingendosi aziende di *check-cashing* e di *debt-collection*. Le licenze di *business* furono ottenute utilizzando le informazioni che provenivano da precedenti furti di identità. Quando ChoicePoint controllò le identità non trovò nulla di sospetto in quanto le informazioni si riferivano a persone realmente esistenti (Otto, Anton, Baumer, 2007).

Otto, Anton e Baumer (2007) continuano spiegando che i controlli di routine di ChoicePoint consistevano nel richiedere ai potenziali clienti che volevano accedere ai loro *database* di dimostrare la propria identità e il motivo per cui si voleva avere accesso. Dopo aver verificato l'identità di un potenziale nuovo cliente, quest'ultimo riceveva una combinazione username/password per accedere al *database*. In questo modo i truffatori hanno portato a termine circa 17,000 ricerche tra i *database* di ChoicePoint, violando il sistema con costi per l'azienda pari a 27,3 milioni di dollari in spese legali, comunicazione alle vittime e ricerca dei colpevoli (Otto, Anton, Baumer, 2007). La FTC indagò sul caso e nel 2006 annunciò che ChoicePoint doveva pagare 15 milioni di dollari (10 milioni in *civil penalties* e 5 per costituire un fondo per risarcire i consumatori) perché aveva violato i termini del *Fair Credit Reporting Act*, condividendo dati di credito personali con utenti non autorizzati e avendo "ingannato" gli utenti facendo credere loro che il loro *database* fosse sicuro (Otto , Anton, Baumer, 2006), (FTC, 2006).

Dopo la violazione, ChoicePoint apportò numerosi cambiamenti alle sue procedure di verifica dell'identità dei consumatori; l'annuncio chiave che fece fu una spiegazione delle condizioni alle quali le informazioni sarebbero state vendute. Oltre a ciò interruppe i rapporti con investigatori privati, esattori e aziende di *check-cashing.* Inoltre, l'azienda istituì controlli più rigidi per l'accesso a codici, password e procedure di disattivazione dell'account.

Crearono anche un ufficio di controllo e rispetto di requisiti e *privacy* che avrebbe monitorato le attività di ChoicePoint e riferito direttamente al Consiglio di Amministrazione, oltre ad un ufficio indipendente per gestire ciò che concerneva la *privacy* (Otto, Anton, Baumer, 2006).

Deborah Platt Majoras, c*hairman* della FTC, dopo il processo disse: "Il messaggio per ChoicePoint e altri dovrebbe essere chiaro: le informazioni personali sui consumatori dovrebbero essere protette dai furti. La sicurezza dei dati è di fondamentale importanza per i consumatori, e proteggerla è una priorità della FTC, come dovrebbe essere per ogni business in America" (FTC, 2006).

3.3 Aziende

Le modalità di trattamento dei dati vengono rese note agli utenti al primo accesso ad un sito web o prima del download di un'APP, tramite *privacy policies* o accordi di licenza. In alcuni casi è strettamente necessario consentire l'accesso ad alcune componenti *hardware* e *software* dello smartphone e ad alcune informazioni con un certo grado di sensibilità: ad esempio, se si è interessati ad un'applicazione che riporta le previsioni metereologiche, è necessario che essa possa accedere alle informazioni relative alla localizzazione dell'utente, al fine di consentire all'APP di fornire le informazioni meteo sul luogo in cui si trova in quel momento l'utente (AGCOM, 2018, p.58).

Proteggere effettivamente la *privacy* sta diventando sempre più difficile con il sempre maggiore utilizzo di Internet. Ci sono però delle azioni che le aziende possono fare per aumentare la fiducia/sicurezza dei consumatori e la trasparenza delle pratiche in cui saranno coinvolti. Esplicitare come i dati dei consumatori verranno utilizzati o se verranno condivisi con terze parti, ad esempio per la creazione di campagne di *marketing* mirate, è certamente un'informazione che dà più consapevolezza ai consumatori. Pubblicare quindi una normativa sulla *privacy* chiara e comprensibile dalla maggior parte degli utenti rende l'azienda più affidabile ai loro occhi e, di conseguenza, si sentiranno più sicuri nella condivisione di informazioni e resteranno fedeli all'azienda (Earp et al., 2005).
Il sito de *Il Post,* un quotidiano italiano *online,* nel quale mi sono imbattuta recentemente, mi è sembrato un buon esempio di questo approccio trasparente. In dettaglio il lettore viene informato di tutti i possibili utilizzi dei dati, viene presentata una lista di tutti i *partners* e per

ciascuno di essi, l'utente può visualizzare le finalità per cui ciascuna azienda utilizza i dati, le loro procedure di *privacy* ed infine, per ciascuna di esse, decidere se accettare o meno di rendere i propri dati disponibili.

3.4 Organismi di controllo

Negli Stati Uniti la FTC è l'ente di controllo deputato al rispetto della privacy; l'agenzia governativa ha il compito di promuovere la tutela dei consumatori e la competitività. La FTC (2013) espone una strategia per cercare di rassicurare i consumatori sulla protezione dei loro dati. Questa strategia si basa su tre punti focali: a) azioni più aggressive di controllo da parte della FTC per far sì che i brokers si attengano al *Fair Credit Reporting Act* (FCRA), che è la legge del governo federale degli Stati Uniti che concerne la *privacy*, promulgata nel 1970; b) ricerche continue, da parte della FTC, per un costante aggiornamento sulle pratiche dell'industria di *brokering*, realtà in continua evoluzione e c) educazione dei *brokers* rispetto alla responsabilità legale che hanno nei confronti dei consumatori, soprattutto i piccoli *brokers* che non sono ancora consci di quali siano i loro obblighi legali. (FTC Internal document, 2013). La FTC (2013) suggerisce che il Congresso degli Stati Uniti dovrebbe rendere esecutiva una legge per dare il permesso ai consumatori di capire quali attività i *brokers* fanno, di dare agli utenti accesso alle informazioni di cui i *brokers* sono in possesso e soprattutto di riconoscere il diritto ai consumatori di vietare lo scambio delle proprie informazioni (*opt-out*). Attualmente infatti solo alcune aziende, tra cui la sopracitata Acxiom, danno il permesso ai consumatori di bloccare la diffusione delle proprie informazioni (totalmente o in parte). Spesso però questo permesso è subordinato al pagamento di una quota di denaro (Deulkar e Gupta, 2018).

Il giornalista Charlie Warzel, in un suo articolo pubblicato su *The New York Times* in Dicembre 2019 riassume alcuni dei punti proposti al Congresso nel *Consumer Online Privacy Rights Act*, un documento che si propone di fornire ai consumatori diritti per la *privacy* dei propri dati e di istituire meccanismi di vigilanza e di controllo:

- Dare ai consumatori la possibilità di vedere, correggere ed eliminare i propri dati, e possibilmente dare loro anche la possibilità di bloccarne la vendita da terze parti.
- Penalizzare le aziende con multe più pesanti in caso di un uso scorretto dei dati.
- Richiedere alle aziende permessi speciali in caso di raccolta di dati sensibili, come ad esempio dati sulla posizione e sulle informazioni fisiche e della personalità.
- Istituire un fondo per la sicurezza dei dati, fondo gestito dal Dipartimento del tesoro.
- Portare le cause sulla *privacy* sotto la gestione della corte federale.

- Pretendere che le aziende mettano a disposizione i propri algoritmi agli enti di controllo (*audit*).

Al momento, il Congresso ha preso tempo perché ritiene che il dibattito sulla *privacy* abbia fatto sì molta strada e che ci sia sicuramente l'urgenza da parte dei cittadini e dei legislatori di chiare linee guida, ma che l'argomento sia troppo complesso per prendere decisioni affrettate (Warzel C., 2019).

L'argomento è indubbiamente importante e delicato ma su questo fronte l'Europa sembra aver fatto passi avanti rispetto agli Stati Uniti. Nell'aprile del 2016, infatti, fu adottato dall'Unione Europea il Regolamento Generale sulla Protezione dei Dati (RGPD) n.2016/679, cioè la normativa europea in materia di protezione dei dati (Saetta B., 2018). Il regolamento è operativo da maggio 2018 e la Commissione Europea si propone come obiettivo quello di rafforzare la protezione dei dati personali di cittadini e residenti dell'UE, sia all'interno che all'esterno dei confini dell'EU.

Successivamente alla pubblicazione del RGPD, lo stato della California ha approvato, il 29 Giugno 2018 il *California Consumer Privacy Act*, la cui entrata in vigore è prevista per il 2020. L'Atto presenta importanti similitudini con il RGPD, ma al tempo stesso anche significative differenze. Ad esempio, mentre in Europa il Regolamento è applicabile ad ogni singolo cittadino, l'Atto Californiano si applica solo ad organizzazioni a scopo di lucro con ben definite caratteristiche (ricavi per almeno 25 milioni/anno di USD, trattamento di dati di almeno 50,000 individui o il cui profitto sia generato almeno al 50% dalla vendita di dati personali) (Privacy.it, 2018).

"Altri stati federali stanno vagliando la possibilità di emanare a breve una normativa di privacy più stringente. Un'onda lunga sembra propagarsi attraverso gli USA dopo che in pochissimi mesi sono arrivati impulsi impossibili da ignorare quali il caso Cambridge Analytica, l'entrata in forza del GDPR (acronimo inglese del RGPD) ed ora, la normativa dello Stato federale che ospita il gotha della rivoluzione digitale" (Privacy.it, 2018).

CAPITOLO 4

IL CASO FACEBOOK-CAMBRIDGE ANALYTICA

Come sopra riportato, il caso *Cambridge Analytica*, ha mostrato come ci sia ancora molta strada da fare per garantire la tutela dei nostri dati. Dopo aver discusso cosa si può fare per evitare di condividere troppe informazioni e cosa le aziende possono fare per tranquillizzare i propri consumatori, vorrei analizzare un caso grave di fuoriuscita di informazioni sensibili.

4.1 I protagonisti

Facebook e *Cambridge Analytica* sono i due agenti principali di questa vicenda. Se tutti noi ben sappiamo cos'è e a cosa serve *Facebook,* penso che non tutti sapessero dell'esistenza di *Cambridge Analytica* prima dello scandalo del 2018. *Cambridge Analytica* (società di consulenza britannica per il *marketing online*) è stata fondata nel 2013 da Stephen K. Bannon, consigliere e stratega di Trump durante la campagna elettorale e poi durante il suo mandato presidenziale, con fondi di un miliardario imprenditore americano molto conservatore, Robert Mercer con l'obiettivo di raccogliere grandi moli di dati sui propri utenti dai *social network* (Menietti E., 2018), ed utilizzarli poi per influenzare, con messaggi mirati e *fake news*, alcune tornate elettorali (vedi Fig. 4.1).

Queste informazioni vengono poi elaborate da modelli e algoritmi per creare dei profili dei consumatori con un approccio simile alla "psicometria", il campo della psicologia che si occupa di misurare le caratteristiche delle personalità, che poi potranno essere arricchiti grazie all'acquisto di ulteriori informazioni dai *data brokers* (Menietti E., 2018).

Durante un'intervista[3] al *The Guardian*, Christopher Wylie, esperto di *data science* che ha lavorato per *Cambridge Analytica* tra il 2013 e il 2014 e principale fonte di informazioni durante questa inchiesta, definisce *Cambridge Analytica* "*a full service propaganda machine"*, in quanto l'azienda ritiene di aver sviluppato dei modelli che non solo riescono a comprendere le preferenze del consumatore e quindi a creare messaggi *ad hoc* per loro, ma riescono anche a far leva sulle emozioni dei consumatori (Wylie C., 2018). Wylie iniziò a collaborare dal 2013 con un'azienda chiamata SCL Group (di cui *Cambridge Analytica* diventerà una sussidiaria), azienda che era specializzata nell'influenzare le elezioni.

[3]Fonte: https://www.theguardian.com/news/2018/mar/17/data-war-whistleblower-christopher-wylie-faceook-nix-bannon-trump

Probabilmente la persona più importante in tutto questo processo fu il Dr. Aleksandr Kogan dell'Università di Cambridge, con il quale *Cambridge Analytica* si mise in contatto. Nel 2014 Kogan aveva ideato un'applicazione chiamata "*thisisyourdigitallife*", un'app che permetteva di produrre profili psicologici e di previsione del proprio comportamento, basandosi sulle attività svolte *online* (Menietti E., 2018), attraverso un test sulla personalità che veniva fatto a chi si iscriveva e che veniva anche remunerato (vedi Fig. 4.1). Gli utenti dovevano accedere all'applicazione usando la modalità "*Facebook login*", grazie alla quale tutte le informazioni necessarie all'installazione della nuova app venivano prese da *Facebook*. Questa applicazione aveva il permesso di raccogliere dal *social network* dati non solo sull'utente che si registrava all'app ma anche sui suoi amici. Circa 270 mila persone installarono l'applicazione dando quindi l'accesso alle proprie informazioni personali (e a quelle degli amici) (Menietti E., 2018) per una quantità enorme di dati. Christopher Wylie aveva ipotizzato che sarebbe stato sufficiente raggiungere un centinaio di migliaia di persone per avere informazioni praticamente su tutta l'America, grazie a questo network di contatti.

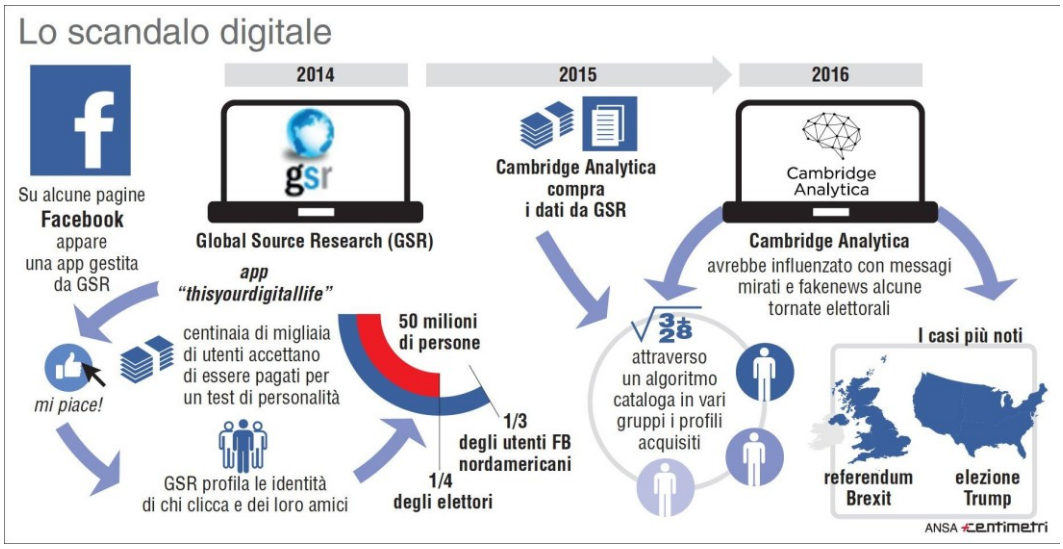

Fig. 4.1: Lo scandalo digitale. Fonte: https://bit.ly/2u6ZjtT, 2018

4.2 Il caso

Si inizia a parlare del caso *Facebook-Cambridge Analytica* a marzo del 2018, dopo la pubblicazione da parte di alcuni importanti giornali, *Guardian* e *New York Times* tra gli altri, di una serie di articoli che dimostravano l'uso scorretto di una grande quantità di dati da parte dei due enti sopracitati, pratiche che avevano molto probabilmente avuto un ruolo durante le elezioni americane del 2017 e anche durante la campagna referendaria per la Brexit.

Con tutte le informazioni in loro possesso, *Cambridge Analytica* aveva la possibilità di costruire un profilo psicologico di tutti gli elettori americani: sapevano a che messaggi gli utenti erano sensibili, composizione, contenuto e tono del messaggio inclusi e grazie a ciò sapevano esattamente quante volte dovevano "colpire" l'utente con un determinato messaggio per fargli cambiare l'opinione su qualcosa (Wylie C., 2018).
Nello specifico furono diffusi una grande quantità di notizie false, post e contenuti contro Hillary Clinton durante tutto il periodo della campagna elettorale, la maggior parte dei quali venivano messi in rete durante i dibattiti televisivi tra Trump e Clinton (Menietti E., 2018).

Il nocciolo dello scandalo e il coinvolgimento di *Facebook* fu l'aver dato la possibilità a "*thisisyourdigitallife"* di impossessarsi delle informazioni degli amici (possibilità comunque indicata nelle infinite pagine di condizioni d'uso di *Facebook*) (Menietti E., 2018). Successivamente questa possibilità venne revocata e Kogan non poté più acquisire informazioni sugli amici, ma questo avvenne troppo tardi. L'app infatti aveva già raccolto dati fino ad arrivare ad avere informazioni di vario tipo su 50 milioni di profili *Facebook*, stima del *New York Times* e del *Guardian* (si veda Menietti E., 2018). I veri problemi però iniziarono quando Kogan decise di condividere tutti questi dati con *Cambridge Analytica*, violando i termini d'uso di *Facebook*, secondo i quali i dati raccolti non possono essere scambiati con terze parti (Menietti E., 2018).
Quella di *Facebook*, infatti, non fu una violazione di dati personali, inteso come *data breach,* bensì una violazione delle normative sulla privacy applicate agli sviluppatori dell'azienda. Kogan aveva detto a *Facebook* che avrebbe usato i dati raccolti per ricerche accademiche e il *social network* non aveva indagato su come venissero effettivamente utilizzate quelle informazioni (Cadwalladr C., 2018).

Oltre al suo ruolo durante la campagna presidenziale di Trump, *Cambridge Analytica* fu al centro di un'altra inchiesta secondo la quale aveva avuto un ruolo anche durante la campagna referendaria per la Brexit. L'azienda aveva utilizzato i dati raccolti per fare propaganda a favore dell'uscita del Regno Unito dall'Unione Europea. La società infatti era in contatto con i maggiori sostenitori del "*Leave*" (Menietti E., 2018).

4.3 Conclusioni del caso

Questa inchiesta partì dal quotidiano britannico *The Guardian* e, anche se non portò a conclusioni definitive e certe, fornì nuovi elementi nel grande dibattito sulle notizie false, sulla propaganda e sulla facilità di diffusione di questi contenuti tramite un uso distorto dei social

network (Menietti E., 2018). L'inchiesta ha messo in luce il fatto che *Facebook* abbia tutt'ora problemi nel controllo dell'uso che viene fatto dei nostri dati e che non abbia strumenti per prevenire un utilizzo distorto delle informazioni: punisce infatti chi non rispetta le regole ma non ha mezzi per evitare che i dati siano scambiati tra terze parti.

Lo scandalo ha decretato il fallimento di *Cambridge Analytica,* mentre dal canto suo Zuckerberg è stato pesantemente multato e ha annunciato che avrebbe preso seri provvedimenti perché casi come quello di *Cambridge Analytica* non si ripresentassero. Tra i correttivi che ha detto di voler apportare, vi sono l'impegno ad effettuare maggiori controlli sulla piattaforma, verificando tutte le applicazioni che hanno accesso a grandi moli di dati e le loro rispettive attività; ad informare, in caso di app sospette, le persone che le hanno usate; a disattivare l'accesso per le applicazioni inutilizzate, ossia se qualcuno non utilizza l'app per un certo periodo di tempo *Facebook* interromperà l'acquisizione delle sue informazioni da parte dell'applicazione; a limitare i dati forniti quando un'app si collega a *Facebook*; a incoraggiare gli utenti ad una gestione più attenta delle app e infine a premiare le persone che segnalano abusi di dati e vulnerabilità (La Rosa A., 2018).

Questo scandalo dimostra come l'acquisizione impropria di dati personali e la loro manipolazione possa avere conseguenze molto pericolose. Tutto ciò ha portato gli attori a capire che su questo tema bisogna essere sempre più informati, consapevoli e attenti e che solo la collaborazione tra tutti gli agenti potrà migliorare l'effetto di questo business sulla nostra quotidianità.

CONCLUSIONE

Decisamente l'aspetto più importante e sensibile nella realtà dei *big data* e dei *data brokers* è relativa all'uso, talvolta improprio, che viene fatto di questi dati, ledendo così la *privacy* dei cittadini. Abbiamo visto che gli Stati Uniti, su cui si è concentrato maggiormente questo lavoro, non hanno di fatto ancora una normativa applicabile a tutti gli stati, alcuni dei quali, come ad esempio la California, si stanno muovendo in autonomia, prendendo spunto dal RGPD.

L'UE ha infatti iniziato a muoversi già da qualche anno creando un regolamento completamente concentrato sulla *privacy* e sul trattamento dei dati personali, con l'obiettivo di dare ai cittadini più controllo sui propri dati, cercando di armonizzare la normativa in modo da renderla omogenea con il resto dell'UE.

Indubbiamente, le azioni dei *brokers* stanno impattando positivamente sia sulle aziende che sui consumatori e stanno cercando di implementare le loro politiche sulla *privacy*, ma in generale si continua a riscontrare mancanza di trasparenza.

L'analisi mette in luce come, anche se crediamo che i dati che diamo a determinati enti siano protetti e non possano essere scambiati, spesso non funziona così.

Come si è visto nei casi *ChoicePoint* e *Facebook-Cambridge Analytica* infatti, anche con la presenza di regolamentazioni per l'utilizzo corretto dei dati, non è ancora possibile controllare tutti gli agenti e c'è sempre la possibilità che avvengano violazioni (*breaches*) dall'esterno o violazioni delle normative di tutela della *privacy.*

Concludendo, sono oggi presenti *online* tantissimi dati (si parla di *zettabyte,* corrispondenti a miliardi di miliardi di *kilobyte*) e per far sì che questi abbiano un impatto positivo sulla nostra vita e sull'economia mondiale è importante che essi vengano analizzati e accorpati con algoritmi sempre più sofisticati, in modo da essere utilizzabili e utili, nel modo giusto e nel rispetto della *privacy*. Avere molti dati e non sapere come usarli, o usarli nel modo sbagliato, può portare a conseguenze dannose per la tutela dei cittadini e pene salate per chi li usa erroneamente. Gli Stati devono impegnarsi nella realizzazione di regolamenti *ad hoc* ma deve esserci interesse da parte di tutti al rispetto di queste normative; noi, in qualità di fonti di dati, dovremmo prestare più attenzione a cosa condividiamo mentre chi raccoglie i nostri dati dovrebbe utilizzarli solo nelle maniere prescritte dalla legge, per poter utilizzare una risorsa, che sembra inesauribile, nel migliore dei modi.

BIBLIOGRAFIA

ACXIOM. *Data, Identity Solution & People-Based Marketing Solutions.* [online]. Disponibile su < https://www.acxiom.com/ > [Data di accesso: 23/06/2020].

AGCOM, 2018. *Big Data.* Interim report nell'ambito dell'indagine conoscitiva di cui alla delibera n. 217/17/CONS. Disponibile su < https://www.agcom.it/documents/10179/10875949/Studio-Ricerca+08-06-2018/c72b5230-354d-444f-9e3f-5467ca450714?version=1.0 > [Data di accesso: 19/06/2020].

BOUNIE, D., DUBUS, A., WAELBROECK, P., 2019. *Collecting and selling personal information: the two faces of data brokers.*

CADWALLADR, C., 2018. "I made Steve Bannon's psychological warfare tool": meet the data war whistleblower, *The Guardian* [online], 18 Marzo. Disponibile su <https://www.theguardian.com/news/2018/mar/17/data-war-whistleblower-christopher-wylie-faceook-nix-bannon-trump > [Data di accesso: 25/06/2020].

CASALI, A., 2019. La Data Monetization As a Service di Olivetti TIM weXplore, *BIGDATA4INNOVATION* [online], 18 Marzo. Disponibile su < https://www.bigdata4innovation.it/internet-of-things/la-data-monetization-as-a-service-di-olivetti-tim-wexplore/#:~:text=%E2%80%9CI%20dati%20sono%20il%20nuovo,grezzo%20non%20pu%C3%B2%20essere%20utilizzato. > [Data di accesso: 25/06/2020].

DEULKAR, D., GUPTA P., 2018. *A Study on Usage of Online Personal Information by Data Broker*. International Research Journal of Engineering and Technology (IRJET). Disponibile su < https://www.academia.edu/38240354/IRJET-_A_Study_on_Usage_of_Online_Personal_Information_by_Data_Brokers > [Data di accesso: 23/06/2020].

EARP, J., et al., 2005. Examining Internet Privacy Policies Within the Context of User Privacy Values. *Engineering Management, IEEE Transactions on*, 52, 227–237.

Effetto GDPR: la California approva la normativa privacy più restrittiva degli USA, 2018. *Privacy.it* [online], 29 Giugno. Disponibile su < https://www.privacy.it/2018/06/29/effetto-gdpr-la-california-approva-la-normativa-privacy-piu-restrittiva-degli-usa/ > [Data di accesso: 25/06/2020].

FEDERAL TRADE COMMISSION, 2013. *What Information Do Data Brokers Have On Consumers, And How Do They Use It.* Washington, D.C.: FTC Disponibile su < https://www.ftc.gov/sites/default/files/documents/public_statements/prepared-statement-federal-trade-commission-entitled-what-information-do-data-brokers-have-consumers/131218databrokerstestimony.pdf > [Data di accesso: 24/06/2020].

FTC, 2006. ChoicePoint Settles Data Security Breach Charges; to Pay $10 Million in Civil Penalties, $5 Million for Consumer Redress. *Federal Trade Commission* [online], 26 Gennaio. Disponibile su < https://www.ftc.gov/news-events/press-releases/2006/01/choicepoint-settles-data-security-breach-charges-pay-10-million > [Data di accesso: 24/06/2020].

KOSKI, H., 2018. *How Do Competition Policy and Data Brokers Shape Product Market Competition?.* Working Papers, ETLA.

LA ROSA, A., 2018. Il caso Facebook-Cambridge Analytica e quello che insegna al mondo della pubblicità digitale. *Programmatic Italia* [online], 23 Marzo. Disponibile su < https://www.programmatic-italia.com/facebook-cambridge-analytica-messaggio-mondo-pubblicita-digitale/ > [Data di accesso: 25/06/2020].

MELISSARI, L., 2018. Il GDPR spiegato per punti. *TPI* [online], 25 Maggio. Disponibile su < https://www.tpi.it/esteri/gdpr-privacy-per-punti-20180525123081/ > [Data di accesso:25/06/2020].

MENIETTI, E., 2018. Il caso Cambridge Analytica, spiegato bene. *Il Post*. [online], 19 Marzo. Disponibile su < http://www.ilpost.it/2018/03/19/facebook-cambridge-analytica/ > [Data di accesso: 24/06/2020].

OTTO, P. N., ANTON, A. I., BAUMER, D. L., 2018. The ChoicePoint Dilemma: How Data Brokers Should Handle the Privacy of Personal Information. *IEEE Security & Privacy Magazine*, 5, 15–23.

RAMIREZ, E., 2014. *Data Brokers. A call for Transparency and Accountability.* Washington D.C.: FTC. Disponibile su < https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf > [Data di accesso: 21/06/2020].

ROSENBERG, M., FRENKEL, S., 2018. Facebook's Role in Data Misuse Sets Off Storms on Two Continents. *The New York Times* [online], 18 Marzo. Disponibile su < https://www.nytimes.com/2018/03/18/us/cambridge-analytica-facebook-privacy-data.html > [Data di accesso: 25/06/2020].

SAETTA, B., 2018. Regolamento generale per la protezione dei dati. *Protezione dati personali* [online], 27 Aprile. Disponibile su < https://protezionedatipersonali.it/regolamento-generale-protezione-dati > [Data di accesso:24/06/2020].

SATTLER, J., 2017. Mikko Hypponen: "Data is the New Oil". *F-Secure* [online], 12 Gennaio. Disponibile su < https://blog.f-secure.com/mikko-hypponen-data-is-the-new-oil/ > [Data di accesso: 24/06/2020].

WARZEL, C., 2019 'Opinion | Will Congress Actually Pass a Privacy Bill?'. *The New York Times* [online], 10 Dicembre. Disponibile su < https://www.nytimes.com/2019/12/10/opinion/congress-privacy-bill.html > [Data di accesso: 24/06/2020].