



**Politecnico
di Torino**

Politecnico di Torino

Corso di Laurea in Ingegneria Gestionale
Percorso Finance
Anno Accademico 2024/2025

Tesi di Laurea Magistrale

Il mercato dei Data Broker: Opportunità, Rischi e Regolamentazione

Relatrice:

Prof.ssa Laura Abrardi

Candidato:

Matteo Raviola



Indice

Introduzione.....	3
1 Introduzione ai data broker.....	5
1.1 Definizione e modello di business.....	6
1.2 Le fonti dei dati.....	8
1.2.1 Documenti pubblici.....	9
1.2.2 Social media.....	10
1.2.3 Attività online.....	11
1.2.4 Abitudini commerciali.....	12
1.3 Elaborazione dei dati raccolti.....	13
1.4 Cessione dei dati.....	17
1.4.1 Modalità di cessione.....	18
1.4.2 Utilizzo dei dati da parte delle imprese a valle.....	19
1.5 Caratteristiche del mercato.....	23
1.6 Competizione e principali attori nel mercato.....	26
1.6.1 Acxiom.....	29
1.6.2 Oracle.....	30
1.6.3 Experian.....	32
1.6.4 Epsilon.....	33
1.6.5 CoreLogic.....	34
2 Rischi.....	36
2.1 Raccolta dei dati online.....	36
2.2 Competition.....	37
2.3 Errori nel profiling.....	40
2.4 Discriminazione di categorie a rischio.....	42
3 Benefici.....	45
3.1 Targhettizzazione.....	46
3.2 Prevenzione delle frodi.....	47
3.3 Verifica dell'identità.....	48
4 Stato attuale delle policy.....	49
4.1 Situazione in Europa.....	49
4.2 Situazione negli USA.....	53



**Politecnico
di Torino**

4.3	Problematiche delle regolamentazioni.....	55
	<i>Conclusioni.....</i>	58
	<i>Bibliografia</i>	63
	<i>Tabella delle figure</i>	67

Introduzione



**Politecnico
di Torino**

Nell'era odierna, quella spesso definita digitale, i dati personali sono diventati una risorsa preziosa e ambita: ogni giorno, infatti, vengono analizzate e vendute miliardi di informazioni riguardanti i consumatori. Le aziende che si occupano di raccogliere, analizzare, manipolare e vendere i dati sono dette data broker. Queste entità operano almeno parzialmente nell'ombra, ottenendo la loro materia prima da molteplici fonti. La relativa invisibilità di questa tipologia di azienda è evidente: nonostante il fatto che le loro attività interessino la quasi totalità della popolazione, sono pochissimi i consumatori che conoscono almeno una delle entità operanti in questo mercato.

L'obiettivo principale dei data broker è la creazione di profili quanto più dettagliati possibili riguardanti gli individui per poi venderli ad enti terzi, ad esempio ad aziende pubblicitarie, istituzioni finanziarie ed enti governativi. Lo studio dei data broker è di fondamentale importanza per comprendere quali sono le implicazioni economiche e sociali della raccolta e dell'uso dei dati personali, settore che, all'interno dell'economia odierna, si rende sempre più indispensabile e onnipresente. È inoltre necessario al fine di comprendere come le normative possano contribuire a rendere positivo il bilanciamento rischi-benefici e a porre un adeguato controllo alla tutela dei diritti dei consumatori, senza frenare eccessivamente la competitività delle aziende.

La finalità di questo elaborato è quella di fornire una panoramica sul mercato dei dati, per poi analizzare rischi e benefici ottenuti grazie alle operazioni delle aziende operanti in questo settore ed esaminare le legislazioni vigenti sull'argomento. Ciò consentirà di trarre delle conclusioni sul bilanciamento tra conseguenze positive e negative dell'azione dei data broker, così come permetterà di ricavare considerazioni finali riguardanti problematiche e sfide dell'attuale situazione regolatoria in alcuni dei principali regolatori a livello mondiale.



**Politecnico
di Torino**

La prima parte di panoramica su questo mercato esplicita la definizione del termine data broker e fornisce un quadro della situazione attuale. Si dichiareranno inoltre le modalità di raccolta, di elaborazione e di cessione dei dati, la materia prima di questa tipologia di azienda. Verranno in seguito nominati i player più rilevanti del settore.

Nella seconda parte verranno trattati in un primo momento i rischi che impattano i consumatori, causati dalle operazioni dei data broker. Alcuni degli errori più rilevanti sono l'erosione della privacy, gli effetti della competition, gli eventuali errori nel profiling e la discriminazione di alcune categorie. In seguito saranno esplicitati i benefici che i data brokers possono fornire al consumatore: la migliorata targettizzazione, la prevenzione di eventuali frodi e la verifica dell'identità

Nell'ultima parte si entrerà nello specifico delle policy che sono state adottate da Europa e dagli Stati Uniti d'America per regolamentare il mercato e tutelare la sicurezza dei dati raccolti. Verranno poi prese in esame le differenze negli approcci dei due regolatori, fondendo ulteriori esempi.

Infine, si valuterà l'esistenza di un bilanciamento tra i rischi e i benefici apportati dal mercato e dalle aziende operanti in esso, così come si effettuerà un'analisi delle problematiche delle legislazioni attualmente vigenti in materia.

1 Introduzione ai data broker



**Politecnico
di Torino**

1.1 Definizione e modello di business

Ogni giorno, i consumatori effettuano molte operazioni online e offline, le quali possono rivelare numerose informazioni su di essi. Che loro ne siano consapevoli o meno, questi dati hanno un valore economico. Le entità con le quali interagiamo li raccolgono, per poi fornirli o venderli a delle aziende dette data brokers. Questo tipo di società opera con modalità molto diverse le une dalle altre, e ciò rende difficile fornire una definizione univoca. Nel 2012 però la US Federal Trade Commission li ha inizialmente descritti come “aziende che raccolgono informazioni sui consumatori, comprese quelle personali, da un'ampia varietà di fonti, allo scopo di rivendere tali dati ai propri clienti, i quali li utilizzano per la verifica dell'identità di un individuo, la differenziazione di un consumatore da un altro, la commercializzazione di prodotti e la prevenzione di frode finanziaria”. Successivamente nel 2014 la definizione fu semplificata in “compagnie che raccolgono informazioni personali per poi cederle o condividerle con altri” [1].

Per quanto riguarda l'Unione Europea, invece, la definizione di data broker non è mai stata introdotta in alcun documento. A seconda dello specifico caso e dell'utilizzo che l'azienda fa dei dati in suo possesso, il Regolamento Generale sulla Protezione dei Dati (GDPR) fornisce le classificazioni “titolare del trattamento”, ovvero l'entità che seleziona mezzi e finalità della gestione delle informazioni, e “responsabile del trattamento”, quindi persona fisica o giuridica



**Politecnico
di Torino**

intitolata della gestione dei dati personali per conto del titolare del trattamento.
[2]

I data broker si specializzano esclusivamente nel commercio dei dati: il loro ruolo all'interno dell'economia delle informazioni è preciso. Essi fungono da intermediari tra le fonti dei dati e le entità che li richiedono. In particolare, questo processo può essere analizzato e diviso in più sottoprocessi: [1]

- ◆ La raccolta di dati, tramite molteplici tipologie di fonti.
- ◆ La trasformazione dei dati in possesso dell'azienda: essi vengono infatti puliti da eventuali duplicati o da errori. Le informazioni provenienti da fonti differenti vengono poi unite, per generare profili unici e quanto più completi possibile, per poter rispondere ad un maggior numero di eventuali richieste. Eventualmente i dataset possono essere successivamente arricchiti da altre informazioni, ad esempio è possibile aggiungere ad un profilo povero di informazioni di tipo creditizio il reddito medio caratteristico della zona nella quale risiede.
- ◆ L'analisi dei dati, che consiste nella clusterizzazione di diverse tipologie di profili, accomunati da una caratteristica comune, quale ad esempio gli interessi o i dati anagrafici. Attraverso tecnologie di machine learning si effettua una predizione analitica, per poter calcolare eventuali rischi di salute, rischio creditizio o anche la probabilità di acquisto di un determinato tipo di prodotto.



**Politecnico
di Torino**

- ◆ L'aggiornamento e la correzione dei dati nel tempo, per mantenerli competitivi e per sottostare a cambiamenti nelle normative vigenti.
- ◆ La monetizzazione dei dati, la quale può avvenire tramite due metodologie distinte. La prima è la vendita: in alcuni casi le informazioni possono essere cedute senza aver effettuato analisi o trasformazioni; infatti, talune aziende preferiscono condurre in autonomia questi procedimenti. In altri casi, vengono venduti dati processati, completi di profilazione e clusterizzazione. La seconda opzione è la concessione da parte dei data broker dell'accesso ai dati in licenza tramite servizi di abbonamento, con un modello di business a sottoscrizione. Ciò consente al data broker di ricevere un flusso di pagamenti costante nel tempo, all'acquirente di avere dati sempre aggiornati e permette anche lo scambio in tempo reale di dati, condizione utile per le campagne pubblicitarie.

1.2 Le fonti dei dati

Il tipo di fonte utilizzato da un'azienda può dipendere dal settore nel quale opera, e quindi della tipologia di dati che necessita. A titolo esemplificativo, una società specializzata nella gestione e vendita di dati di tipo finanziario sarà qualificata nella ricerca tramite alcuni canali piuttosto che altri.

Le fonti possono essere distinte in primarie e secondarie: si parla di fonti primarie quando i dati vengono raccolti direttamente dalla loro fonte, come ad



**Politecnico
di Torino**

esempio tramite documenti pubblici. Quando la fonte viene invece denominata secondaria, si tratta di dati acquistati da un'altra entità di raccolta e vendita di informazioni: spesso un altro data broker. [3]

1.2.1 Documenti pubblici

Vi sono differenti tipologie di documenti pubblici, e ognuno di questi fornisce dati che possono essere utilizzati per diversi scopi. Ad esempio, i registri immobiliari forniscono informazioni riguardo alle proprietà immobiliari di un soggetto, che sia un consumatore o un'azienda, ma anche riguardo eventuali ipoteche e transazioni passate. Sempre tramite i registri pubblici è possibile reperire i dati fiscali di un cittadino, nei quali rientrano imposte pagate, dichiarazione dei redditi, eventuali debiti fiscali e redditi. Sono ottenibili, inoltre, i documenti che legano il consumatore ai tribunali, come procedimenti legali e casi giudiziari, e chiaramente le informazioni demografiche.

La maggior parte di queste informazioni possono essere ricavate da qualsiasi privato cittadino, ma i canali per fare ciò sono tortuosi e sarebbe molto complicato ottenere alcuni permessi e pagare per eventuali portali. In aggiunta, sono rilevanti i volumi di informazioni ricavati tramite le partnership commerciali con i fornitori di dati che hanno accordi con gli enti governativi. [4] [5]



**Politecnico
di Torino**

1.2.2 Social media

Una fonte ormai sempre più rilevante per i data broker è l'attività sui social media. Vi sono più modalità attraverso le quali le informazioni personali possono essere raccolte. Una di queste è la pubblicazione volontaria: i social network si basano sulla condivisione di foto, pareri e testi, e tutto ciò può essere raccolto e analizzato. Un "Mi piace" ad una pagina, così come la condivisione di una posizione o una foto in un luogo specifico, sono elementi utili per la raccolta di informazioni tramite scraping, insieme di tecniche che permettono l'automatizzazione della ricerca dei dati dei consumatori.

Sono molto utilizzate anche le API (Application Programming Interface): sono numerosi i social media che consentono agli sviluppatori o a terze parti di accedere ai dati degli utenti, richiedendone il loro consenso. Questo viene fornito per velocizzare alcune procedure, come la creazione di un account su un sito, ma permette la condivisione dei nostri dati.

Gli stessi social media, poi, collaborano con aziende di analisi: i primi raccolgono grandi volumi di informazioni sui loro utenti, dalle quali verranno estratti i dati più rilevanti e utili, tramite data mining e machine learning.



**Politecnico
di Torino**

Tutte queste metodologie consentono alle aziende che acquistano i dati dai data broker di ottenere un profilo dettagliato del consumatore, a collocarlo nel giusto segmento di mercato e quindi ad attuare strategie di marketing personalizzate ed altamente efficaci. Questo non è però l'unico utilizzo che si può fare delle informazioni degli utenti: ciò si è dimostrato vero nel caso di Cambridge Analytica. Questa azienda di analisi di dati, infatti, è finita sotto i riflettori nel 2018, quando si è scoperto che avesse ottenuto i dati di circa 87 milioni di utenti di Facebook sottoponendo un test di personalità a 270000 persone: la piattaforma consentiva infatti la condivisione dei dati di tutti i profili collegati tra loro. [6] Ciò è stato possibile a causa di regolamentazioni legislative e interne alla compagnia di social network troppo poco strette.

1.2.3 Attività online

La semplice navigazione sul web può fornire molte informazioni su un utente: infatti, lo stesso motore di ricerca può tracciare i siti visitati e le ricerche online, mentre i singoli siti possono, tramite l'utilizzo dei cookie, memorizzare le preferenze dell'utente, nonché raccogliere dati sul comportamento dello stesso sulla pagina web.



**Politecnico
di Torino**

Da non trascurare sono i servizi che i consumatori utilizzano per rendersi alcuni compiti più agevoli: uno di questi può essere un'applicazione di geolocalizzazione: rendendo disponibile la nostra posizione, stiamo volontariamente fornendo una quantità di informazioni enorme, a partire dai luoghi da noi frequentati più assiduamente fino ad arrivare ai nostri orari e al nostro stile di vita.

1.2.4 Abitudini commerciali

Nonostante il tempo speso online dal consumatore medio aumenti di anno in anno, anche le azioni fisiche consentono agli aggregatori di dati di svolgere la loro funzione. Ad esempio, qualsiasi acquisto noi effettuiamo senza l'utilizzo dei contanti viene registrato, e possiamo essere profilati sotto il punto di vista del reddito e delle abitudini. Ciò è valido anche per i programmi fedeltà e premi: associando una tessera a punti ad un consumatore, ogni suo acquisto può essere memorizzato. Alcuni di questi, poi, hanno una particolare rilevanza: possono essere ricavate informazioni assicurative, quali ad esempio dati sulle polizze e sui sinistri, ma anche creditizie, come debiti e punteggi di credito.

Anche i servizi forniti dalle compagnie telefoniche sono una fonte rilevante di informazioni: esse per loro natura sono in possesso di dati di miliardi di consumatori. I numeri di telefono con i quali gli utenti



**Politecnico
di Torino**

interagiscono, l'orario delle chiamate e anche la loro durata sono alcune delle informazioni a disposizione delle aziende operanti in questo settore. [7] Tutte queste tipologie di dati sono state rese disponibili al governo degli Stati Uniti, come rivelato dal caso Snowden. [8]

1.3 Elaborazione dei dati raccolti

I dati raccolti o acquistati devono essere memorizzati, e ciò può avvenire con differenti modalità:

➤ **Server locali**

Sono data center aziendali. Questa modalità di conservazione dei dati ne consente un controllo più stretto rispetto agli altri: ogni aspetto della memorizzazione e del mantenimento è a carico del data broker. L'hardware è infatti di sua proprietà, e con esso la responsabilità della sicurezza logica (protezione dei dati tramite backup e crittografia e delle reti, tramite firewall e VPN) e di quella fisica (sistemi di raffreddamento, di videosorveglianza, guardie di sicurezza). I costi iniziali di setup sono alti, e la scalabilità bassa.

➤ **Server esterni**



**Politecnico
di Torino**

Sono data center di un'altra compagnia: viene affittato lo spazio fisico e le risorse di elaborazione. La memorizzazione avviene all'interno di hardware di proprietà del data broker, che è responsabile della sua gestione e manutenzione. Il data center fornisce invece alimentazione, raffreddamento e sicurezza fisica. La sicurezza logica è gestita dal data broker. Fornisce quindi un minore livello di controllo su alcuni aspetti della memorizzazione dei dati.

➤ Cloud

Simile ai server esterni, il cloud è più flessibile e scalabile: l'hardware si trova in data center con enormi capacità di memorizzazione e calcolo, e la sua manutenzione è infatti responsabilità del provider, così come la sicurezza, sia logica che fisica. L'accesso ai dati può avvenire tramite API oppure tramite pannelli di controllo.

Questa modalità di memorizzazione è ampiamente la più utilizzata dai data broker.

➤ Soluzioni miste

Un approccio ibrido consiste nell'applicazione simultanea di una memorizzazione cloud e di una in server locali. Questi serviranno per memorizzare sia i dati utilizzati più di frequente sia quelli che per regolamentazione non possono essere memorizzati in cloud. Questa



**Politecnico
di Torino**

soluzione fornisce un compromesso tra sicurezza e scalabilità, compromettendo però la facilità di gestione.

Dopo aver raccolto o ottenuto i dati, questi vengono manipolati. Il mercato della compravendita di questi, infatti, può essere diviso in

- **Dati non strutturati**

Questa tipologia di informazioni non è organizzata in un formato predefinito o in un database strutturato: non seguendo un modello rigido, possono essere memorizzati in diversi modi. Questi dati possono essere ad esempio testi liberi (e-mail, post sui social, articoli), documenti o media. Lo svantaggio di questo tipo di memorizzazione è la necessità di utilizzo di tecniche di elaborazione avanzate per l'analisi.

- **Dati strutturati**

Sono informazioni organizzate seguendo formati predefiniti, quali database relazionali e tabelle. Seguono uno schema coerente e accessibile, il quale permette una rapida analisi e manipolazione, tramite linguaggi di query.

- **Dati con struttura personalizzata**



**Politecnico
di Torino**

In questa tipologia di memorizzazione, le informazioni vengono organizzate in modo specifico, per soddisfare le esigenze particolari di un'azienda o di un progetto.

I dati non strutturati sono la fetta più ampia all'interno del mercato, plausibilmente oltre il 50% del suo valore totale. [9]

A seconda della tipologia di dati, questi subiscono svariati passaggi prima di poter essere venduti. Se sono di tipo non strutturato, vengono prima organizzati e standardizzati. Una volta resi quanto più strutturati possibile, vengono rimossi i duplicati e corretti gli errori.

Quello evidenziato nella Figura 1 è il risultato di un processo fondamentale, ovvero quello della creazione di profili di singoli individui. Questi profili, che includono informazioni finanziarie, demografiche e comportamentali, vengono resi unici tramite codici univoci interni e proprietari del data broker, [7] e le informazioni provenienti da fonti diverse vengono associate al relativo consumatore. Acxiom, ad esempio, fornisce fino a 3000 attributi per ciascuno di essi, combinando molteplici tipologie di dati. [7]



Politecnico di Torino

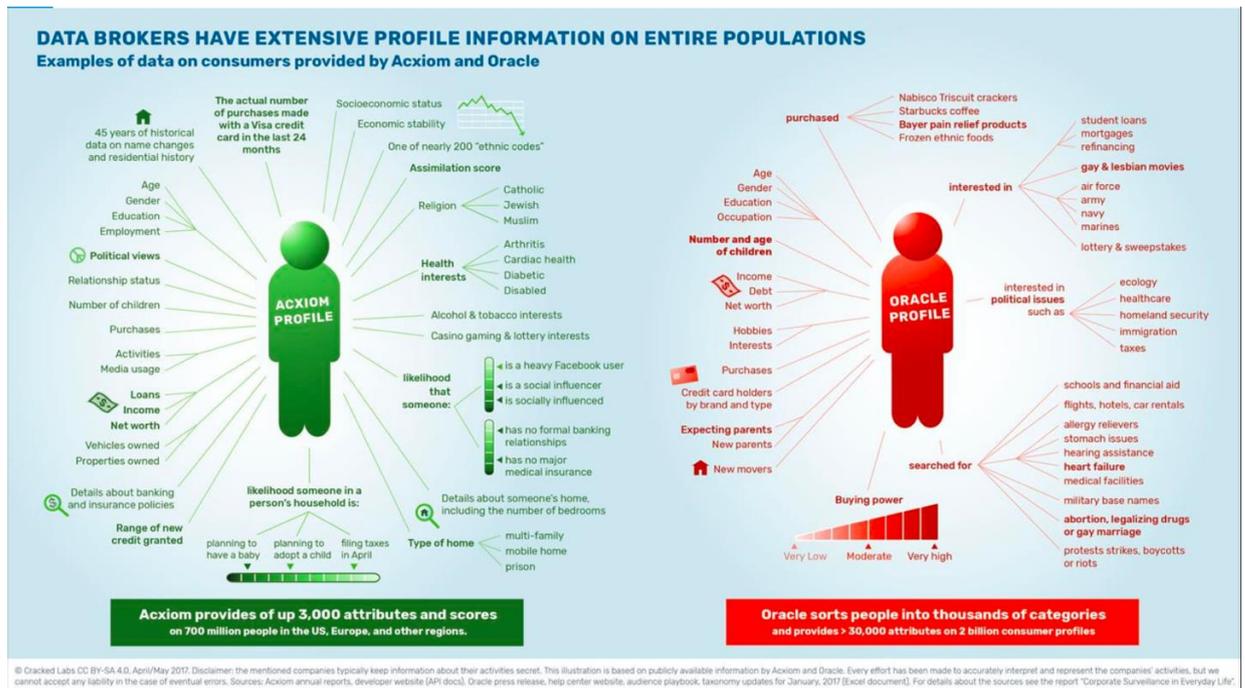


Figura 1: Profilazione dei consumatori [7]

Questi profili sono poi raggruppati in categorie con caratteristiche comuni, quali età, interessi, abitudini, reddito, in modo da essere più facilmente targettizzabili a seconda delle necessità del cliente.

Vengono poi analizzati, per ottenere pattern e tendenze, mediante l'utilizzo di analisi avanzate.

1.4 Cessione dei dati



**Politecnico
di Torino**

1.4.1 Modalità di cessione

Per monetizzare i dataset in loro possesso, i data broker hanno la possibilità di scegliere due modalità differenti. La prima è la concessione in licenza: il cliente in questo caso paga una sottoscrizione periodica per poter accedere al dataset in maniera continuativa. Questa tipologia di servizio tipicamente include aggiornamenti e correzioni del database, e consente di risolvere il problema della perdita di valore dei dati. Il data broker avrà però l'onere di mantenere il dataset aggiornato per non perdere i clienti che pagano la sottoscrizione. In questo caso i ricavi dell'azienda di data broker sono solitamente prevedibili.

La seconda modalità di cessione dei dati è la vendita: consiste nella singola fornitura dell'intero dataset richiesto, senza ulteriori aggiornamenti. È una soluzione utile, ad esempio, per campagne di marketing che non si protraggono nel tempo o per ricerche. Il fatturato in questo caso dipende dal numero di singole vendite che si riescono ad effettuare. I dataset possono perdere di valore nel tempo, a causa dell'obsolescenza dei dati, e il data broker per rimanere competitivo e attrarre nuovi clienti deve aggiornare continuamente il database.

La scelta tra questi due modelli di business dipende dalla nicchia di mercato che occupa il data broker: ad esempio Experian utilizza principalmente vendite singole, [10] per poter fornire report creditizi, mentre invece Acxiom utilizza maggiormente le sottoscrizioni. [11]



**Politecnico
di Torino**

1.4.2 Utilizzo dei dati da parte delle imprese a valle

Nella Figura 2 è riportata una delle possibili classificazioni dei dati che vengono ceduti dai data broker, sotto il punto di vista dell'utilizzo finale che il cliente ne farà. Queste categorie principali sono: dati con finalità di marketing, di mitigazione del rischio e personali.



Figura 2: Scopi dei dati venduti [1]



**Politecnico
di Torino**

Le aziende che acquistano i dati li integrano all'interno dei CRM, customer relationship management, sistemi di gestione interni. Questi consentono di migliorare la profilazione degli utenti tramite l'integrazione dei dati già in possesso dell'azienda, e quindi consentire una superiore analisi predittiva.

I dataset acquistati verranno poi utilizzati per una grande varietà di possibili scopi, dipendenti dalla tipologia di entità che li ha acquistati. Alcuni di questi sono:

- **Analisi del mercato**

Le società possono comprendere i comportamenti dei consumatori e la variazione delle loro abitudini. Ciò migliora la conoscenza delle tendenze di mercato, consente un'analisi della situazione delle aziende rivali e può aiutare nell'identificazione di nuove opportunità di business.

- **Marketing**

La creazione di campagne pubblicitarie mirate e personalizzate passa per forza di cose dal possesso dei dati di quanti più consumatori possibile. Ciò aumenta sensibilmente l'efficacia del budget di marketing delle aziende.



**Politecnico
di Torino**

Lo scopo di questi dati può essere quello di creare clusterizzazioni di clienti che condividono caratteristiche simili, per poter agire su diversi aspetti del loro comportamento. È possibile quindi personalizzare le campagne pubblicitarie, in base a molteplici fattori, quali la posizione geografica, la risposta a campagne passate, e la situazione reddituale del soggetto. [1]

- **Mitigazione del rischio**

Le aziende che acquistano dati possono utilizzarli per diminuire varie tipologie di rischi. Essi possono essere rischi interni o esterni. Lo scopo di mitigare i primi è quello di fornire un servizio ai clienti, quali ad esempio la prevenzione di furti di identità e di frode. La riduzione dei rischi interni invece permette di diminuire le perdite di denaro dovute alle operazioni aziendali. Infatti, tramite l'analisi dei dati è possibile mitigare rischi di tipo operativo, di supply chain e di mercato. Un esempio sono le informazioni utilizzate dalle istituzioni finanziarie: queste acquistano i dati riguardanti reddito e affidabilità dei clienti per prendere decisioni riguardo a prestiti e altre operazioni finanziarie. Il settore che più degli altri contribuisce al mercato dei data broker è proprio quello degli istituti finanziari, ed esso costituisce la fetta più ampia dei ricavi di questa industria. [9]

- **Assistenza clienti**



**Politecnico
di Torino**

I dati utili per migliorare la qualità del servizio clienti di un'azienda sono ad esempio i dati demografici, quelli comportamentali (canali di comunicazione e fasce orarie predilette), dati storici (identificazione dei clienti fidelizzati, conoscere gli ultimi acquisti), ma anche feedback e opinioni pubblicati su internet.

- Rivendita

I data brokers possono acquistare i dati per implementare quelli in loro possesso: lo scambio è infatti una delle principali fonti di dati. [12] Ciò è possibile perchè all'interno dei contratti di acquisto di informazioni, spesso vi sono clausole che permettono ai data broker di ottenerne la proprietà. [7]

Inoltre, i data broker spesso vengono in possesso di informazioni che esulano da quelle utili al loro modello di business. Per poter monetizzare anche questi dati, una delle opzioni disponibili è quella di rivenderli ad altre aziende dello stesso settore.

- Scopi governativi

Anche i governi acquistano dai data broker i dati dei cittadini, per scopi di intelligence: ad esempio, possono risultare utili per fare prevenzione del crimine, analizzando transazioni e dati di geolocalizzazione, ma anche per gestire eventuali emergenze, quali



**Politecnico
di Torino**

pandemie o disastri naturali, ottenendo informazioni in tempo reale e ottimizzando l'allocazione delle risorse. [13]

1.5 Caratteristiche del mercato

Quello dei data broker è un mercato a rapida crescita, e gioca un ruolo fondamentale nell'aiutare altri business a capitalizzare il potenziale dell'utilizzo dei dati dei consumatori.

Nella Figura 3 si possono vedere le stime della dimensione del mercato e della sua evoluzione nel tempo: nel 2012 il fatturato veniva stimato essere pari a 150 milioni di euro [14] e a 200 nel 2016 [15], per poi raggiungere quota 232 nel 2019 [16]. Dal 2012 al 2019 il tasso composto di crescita annuale è stato di poco inferiore al 6,5%. Il CAGR è poi diminuito di 3 punti percentuali nel periodo dei successivi 4 anni, arrivando ad un valore di 264 miliardi nel 2023.



**Politecnico
di Torino**

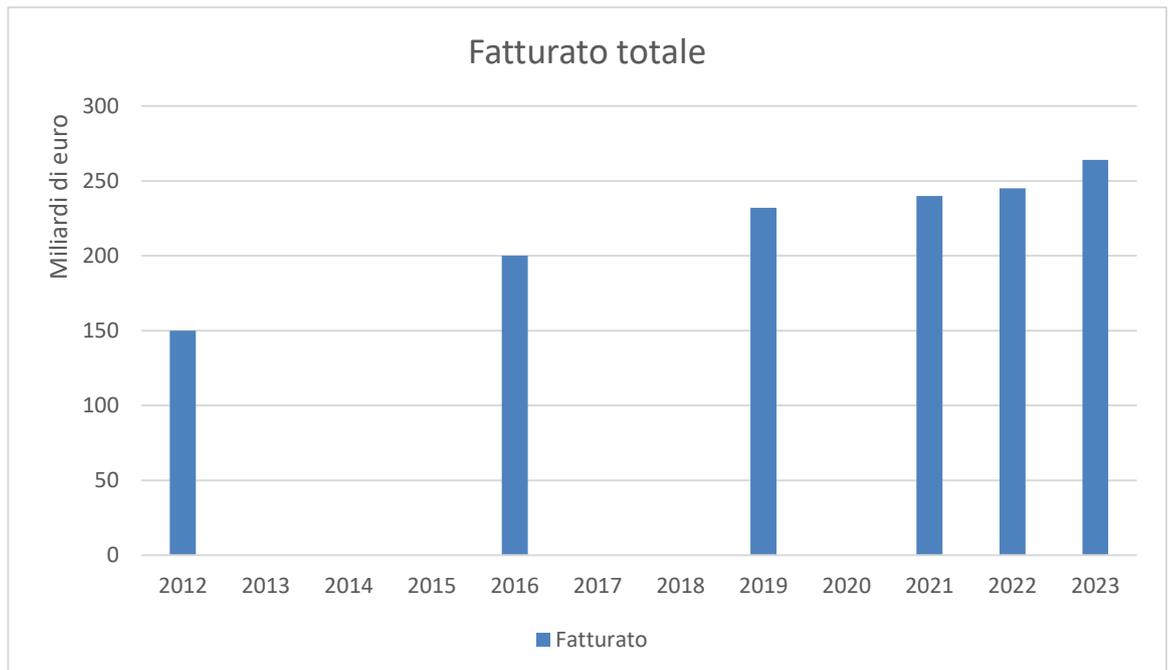


Figura 3: Crescita del mercato dei data broker [14] [15] [16] [17] [18] [19]

Per il fatturato del mercato dei data broker, la crescita futura prevista dalle analisi più recenti è pari all'8%, e arriverà a valere mezzo trilione di euro nel 2028. [17] Tuttavia, si tratta di stime: è infatti complicato ottenere un valore univoco della dimensione e del tasso di crescita del mercato, data l'ampia varietà di operazioni che possono essere comprese al suo interno: ad esempio, le aziende big tech non vengono solitamente incluse, dato che raccolgono i dati direttamente da utenti che hanno fornito il loro consenso. Nonostante ciò, il rapporto tra i social media e i data broker appare fumoso e non definito, e sono



**Politecnico
di Torino**

numerosi i casi conosciuti di rapporti stretti tra i due settori, come ad esempio quello tra Facebook e DataLogix. [20]

Quello della raccolta e vendita di dati è un settore complesso, che si compone di più livelli di data brokers, tra i quali vi sono complesse connessioni. Infatti, le informazioni vengono raccolte da molteplici fonti, e quando vengono acquistate da un data broker, questo le combina a quelle in suo possesso, per poi spesso rivenderli anche ad altri competitors. Ciò rende complicato risalire alla fonte originale di un singolo attributo di un singolo consumatore, e infatti la maggior parte dei dati in possesso dei data broker proviene da altre aziende del mercato, piuttosto che da fonti primarie. [21]

Per quanto riguarda la suddivisione geografica del fatturato di questo settore, il nord America si aggiudica la porzione più ampia, 115 miliardi di euro, pari a circa il 40% del valore totale del mercato dei data broker. [9] In seconda posizione troviamo l'Europa: come riportato nella Figura 4, la Germania si aggiudica la leadership del continente, seguita da Regno Unito (il mercato a più rapida espansione nella zona), Russia, Francia e Italia. Il valore totale del mercato europeo ammonta a circa 80 miliardi di dollari. [22] [23]



**Politecnico
di Torino**

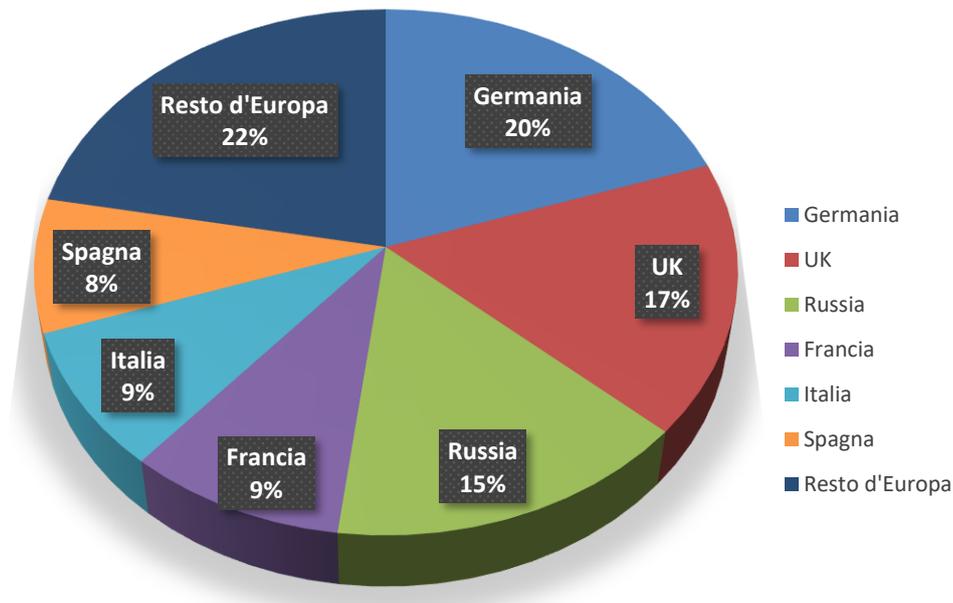


Figura 4: Suddivisione del mercato in Europa [22]

Il valore del mercato dei data broker del continente asiatico, invece, è quello che si prevede crescerà di più nel prossimo decennio, con Cina e India destinate ad aumentare sempre più le loro quote di mercato. [9]

1.6 Competizione e principali attori nel mercato

Per conoscere le dinamiche di competizione e di vantaggio competitivo dei data broker è possibile analizzare le 5 forze di Porter:



**Politecnico
di Torino**

1. Minaccia di nuovi concorrenti

Le barriere all'entrata sono relativamente alte. Ciò è dovuto agli investimenti che devono essere effettuati per poter competere in tecnologia e infrastrutture di raccolta, analisi e conservazione di enormi quantità di dati. Anche le stringenti norme sulla privacy pongono una barriera all'ingresso, scoraggiando ulteriori aziende ad entrare nel mercato. Inoltre, se un nuovo concorrente trovasse una nicchia e riuscisse a competere, sarebbe probabile un'acquisizione da parte delle società più grandi: le incorporazioni sono infatti un importante driver della loro crescita. [24] [25]

2. Potere contrattuale dei fornitori

Il potere contrattuale dei fornitori è ridotto: i data broker possono acquistare o ottenere i dati da molteplici fonti, e anche scambiarseli tra loro. Ciò non avviene nel caso di dati settoriali e particolarmente specifici: in questa casistica il potere contrattuale dei fornitori è più alto [26]. Per quanto riguarda invece i fornitori di servizi di memorizzazione di dati, in particolare di cloud, la loro posizione dominante può consentire loro di esercitare un certo potere. [27]

3. Potere contrattuale degli acquirenti



**Politecnico
di Torino**

Le aziende che acquistano i dati possono avere un potere contrattuale non trascurabile, ma funzione della loro dimensione: nel caso di società molto grandi, i volumi di dati acquistati sono rilevanti, mentre per quanto riguarda le PMI il potere contrattuale è sbilanciato in favore dei data broker. La situazione è stata presa in considerazione nel Regolamento EU noto come Data Act: si è tentato di diminuire lo squilibrio di potere nei contratti di acquisto di dati da parte delle piccole e medie imprese [28].

4. Minaccia di servizi sostitutivi

Il sopravvento di nuove tecnologie potrebbe mettere a rischio la posizione dei data broker: questi possono essere definiti infatti degli intermediari, e le grandi aziende potrebbero sviluppare la capacità interna di raccolta e analisi di dati, anche grazie all'intelligenza artificiale. In seguito a ciò potrebbe nascere un commercio dei dati dei consumatori tra aziende di diversi settori. Ad oggi i data broker rimangono fondamentali per le campagne marketing e di profilazione della maggior parte delle compagnie. [29]

5. Rivalità tra competitor esistenti

Pur essendo un mercato concentrato, le aziende più grandi spesso sono in possesso degli stessi dati, a causa dell'utilizzo delle stesse fonti e anche della compravendita delle informazioni sui consumatori. Ciò comporta un basso costo di switching da parte delle aziende clienti. D'altra parte, vi sono alcuni data broker che si specializzano nella raccolta dei dati di una certa



**Politecnico
di Torino**

tipologia, ad esempio quelli reddituali o immobiliari. Questo comporta una competizione ridotta per questo tipo di aziende. [12]

L'analisi delle forze di Porter rivela un settore complesso e competitivo. I player esistenti non temono l'entrata di nuovi competitor al suo interno, ma la concorrenza degli altri data broker. Le economie di scala sono rilevanti: i costi fissi possono essere facilmente diluiti se il volume di dati processato e venduto è di grandi dimensioni. Le aziende più importanti sono le seguenti:

1.6.1 Acxiom

Fondata nel 1969 con il nome di Demographics, Inc. in una prima fase era specializzata nella gestione dei dati per il marketing diretto.

Si è rapidamente espansa nel mercato della raccolta di informazioni, sbarcando in Europa nel 2004, e tramite numerose acquisizioni ha aumentato sempre di più la sua rilevanza. Nel 2012 era in possesso del database sui consumatori più vasto al mondo, e due anni più tardi ha acquistato LiveRamp, un'importante piattaforma di gestione di dati. La parte del business che si occupava di marketing diretto è stata venduta nel 2018 per 2,3 miliardi di dollari, e in seguito ha cambiato il suo nome in LiveRamp. [30] [31]



**Politecnico
di Torino**

Ad oggi è una multinazionale con una capitalizzazione di mercato pari a 2 miliardi di dollari, con sedi negli Stati Uniti, in Europa e in Asia, e impiega più di 2000 persone. Ha registrato un fatturato di 660 milioni di dollari nel 2023 e utili netti pari a 12 milioni di dollari. La maggior parte del fatturato, pari al 78%, deriva dagli abbonamenti, metodologia principale attraverso la quale il data broker fornisce i dati ai propri clienti. Possiede i dati di circa 700 milioni di consumatori, e fornisce fino a 12000 attributi per ogni singolo consumatore nei loro database [32]. Si tratta infatti di un data broker che raccoglie un ampio spettro di tipologia di dati, senza specializzarsi in una tipologia specifica.

1.6.2 Oracle

È una delle principali aziende tecnologiche a livello mondiale. Fondata nel 1977 con il nome Software Development Laboratories, nel 1982 venne rinominata Oracle Systems Corporation.

L'obiettivo dell'azienda era la creazione di un sistema di gestione di database relazionali. Questo venne raggiunto nel 1979, e il nome commerciale del software diede il nome all'azienda. A partire dal 1983 iniziò la sua espansione internazionale, aprendo uffici in Europa e Asia. Tre anni più tardi si quotò in borsa, continuando ad aggiornare il prodotto di punta, Oracle. Nel corso degli anni seguenti l'azienda



**Politecnico
di Torino**

effettuò numerose acquisizioni per diversificare il suo portafoglio di servizi software. La più rilevante fu sicuramente l'incorporazione nel 2009 di Sun Microsystems, per un valore di 7,4 miliardi di dollari. nel frattempo, a partire dal 2007 l'azienda entrò nel mercato del cloud computing, ad oggi tecnologia fondamentale per Oracle, e nel 2019 strinse una partnership con la rivale Microsoft, collegando i rispettivi cloud. La società entrò inoltre nel settore dell'intelligenza artificiale, effettuando numerose acquisizioni dal 2021 in poi. [33]

Il ramo dell'azienda che opera nel mercato dei data broker è denominato Oracle Advertising, ed è stato sensibilmente accresciuto grazie in particolare a due incorporazioni strategiche avvenute nel 2014, quella di BlueKai (piattaforma di gestione di dati) e quella di Datalogix, azienda di raccolta dati [34]. Grazie a questa espansione, la società possiede i dati di 5 miliardi di persone, più del 60% dell'intera popolazione mondiale, e di circa il 90% di quella statunitense. Ciò rende Oracle la società in possesso del database di informazioni sui consumatori più vasto al mondo. Questo dato è stato citato nel 2016 da uno dei fondatori di Oracle, Larry Ellison, e si riferisce ad una causa intentata ai danni dell'azienda. È però un numero eccessivamente elevato, e la reale dimensione del database in possesso della multinazionale non è pubblico, ma è stimato essere pari a circa due miliardi di utenti unici [7]. Il dataset è stato ottenuto da tre fonti principali: l'acquisizione di Datalogix, di BlueKai e di AddThis



**Politecnico
di Torino**

(piattaforma di data mining proprietaria). Oracle non si specializza nella raccolta di dati di una tipologia specifica, e non condivide informazioni più approfondite a riguardo.

Ad oggi, Oracle Corporation registra ricavi annuali pari a 50 miliardi di dollari, con un utile netto pari a 8,5 miliardi di dollari, ma i risultati economici del ramo della gestione e vendita di dati non vengono riportati separatamente: essi sono infatti compresi in quelli della sezione “Cloud services and license support”, la quale ammonta al 71% dei ricavi totali [35].

1.6.3 Experian

Originariamente nota con il nome di Credit Data Corporation, è stata acquistata e poi rivenduta da tre aziende dal 1968 al 2006, rispettivamente dalla americana TRW, operante in numerosi settori, successivamente dalla Bain Capital, compagnia di investimento, e infine dalla GUS, che ai tempi era la più grande azienda inglese che forniva servizi riguardanti il credito. Il 2006 è stato l’anno della quotazione sulla borsa di Londra [36].

È un data broker specializzato nel credit reporting: raccoglie quindi dati sulla situazione finanziaria dei consumatori, in particolare qualsiasi informazione utile per gestire il rischio del credito, per prevenire frodi e



**Politecnico
di Torino**

per targettizzare con maggiore precisione i consumatori. Questo tipo di dataset è ormai fondamentale per banche e istituti finanziari di varia natura con lo scopo di valutare l'affidabilità creditizia dei clienti, per aziende intenzionate ad avviare campagne di marketing ma anche a proprietari di immobili, per trovare potenziali affittuari.

Experian è una delle principali società globali focalizzata nel credit reporting, con una capitalizzazione di mercato di 36 miliardi di dollari, e un fatturato pari a 7 miliardi nel 2023. Il 66% di questo è originato negli Stati Uniti, l'America Latina contribuisce per il 16% mentre il Regno Unito per il 12%. Gli utili netti ammontano a 1,3 miliardi di dollari. [37]
[38]

Negli anni Experian ha raccolto dati su 2,3 miliardi di persone, e possiede informazioni di carattere creditizio su più di 900 milioni di consumatori. [7]

1.6.4 Epsilon

È stata fondata nel 1969, per poi venire acquisita dalla multinazionale francese Publicis Group nel 2019 per un valore di circa 4 miliardi di euro. Epsilon è uno dei data broker più rilevante a livello globale: è in possesso, infatti, di dati di 250 milioni di persone residenti negli Stati Uniti [39], pari al 75% della popolazione. Il suo fatturato è stimato



**Politecnico
di Torino**

intorno agli 1,7 miliardi di dollari: esso non viene infatti esplicitamente dichiarato dalla società madre, la quale nel 2023 ha registrato ricavi pari a 4,5 miliardi di euro. [40]

Si tratta di un data broker che non si specializza nella raccolta di una tipologia particolare di informazioni: TargetSource, il database proprietario di Epsilon, contiene infatti dati che spaziano da quelli di tipo demografico a quelli transazionali e finanziari. Una delle fonti principali utilizzate da Epsilon è ShoppersVoice, la più grande azienda americana di ricerche di mercato. [41]

1.6.5 CoreLogic

Nel 1991 il ramo di raccolta dati immobiliari di TRW avviò una partnership con Reed Elsevier, multinazionale di analisi dati e pubblicità. Cinque anni più tardi, nel 1996, questa parte di business venne divisa dall'azienda madre in seguito ad uno spin off. Questa nuova compagnia venne nominata Experian, oggi competitor di CoreLogic. Acquisita da un'azienda immobiliare, Experian venne accorpata con CoreLogic Systems nel 2007. Nel 2010 CoreLogic fu separata dalla compagnia principale, e da allora iniziò una lunga serie di acquisizioni per ampliare l'offerta di servizi. Nel 2021 è stata acquisita da una azienda di private equity, e ad oggi CoreLogic è un data broker specializzato in



**Politecnico
di Torino**

informazioni immobiliari, creditizie, di rischi e disastri ambientali. È quindi fornitore di analisi del rischio creditizio ma anche di dati geospaziali per il settore immobiliare e per quello delle assicurazioni, e fornisce inoltre previsioni dell'impatto che il cambiamento climatico avrà su un determinato immobile. [41] L'ultimo fatturato noto è datato 2020, e ammontava a 1,8 miliardi di dollari [42]. Il database di questa azienda contiene 795 milioni di transazioni di proprietà immobiliari, e il 99% delle proprietà degli Stati Uniti. [1]

Ciò che accomuna tutte le aziende operanti in questo settore, in particolare quelle che dispongono di grandi quantità di dati, sono i problemi di sicurezza e di privacy. Sono infatti numerosi i casi di data breach, ovvero di furti di informazioni: essere totalmente invulnerabili agli attacchi informatici è molto difficile, e quando uno di questi va a segno, le informazioni sensibili di milioni di persone vengono pubblicate online o viene richiesto un riscatto. Altra caratteristica del settore è l'elevato numero di class action intentate ai danni delle aziende: spesso i consumatori non sono soddisfatti della modalità di gestione ma soprattutto di raccolta dei loro dati personali, e avviano cause milionarie contro le società che secondo il loro punto di vista violano la loro privacy.



**Politecnico
di Torino**

2 Rischi

Il modello di business caratteristico dei data broker comporta implicitamente dei rischi per i consumatori, che sono i soggetti ai quali i dati vengono collegati. Le implicazioni che questi rischi comportano possono spaziare in diversi ambiti, e compromettere significativamente alcuni contesti della vita privata dei consumatori. Essi spesso non sono nemmeno a conoscenza della fonte della loro problematica.

2.1 Raccolta dei dati online

Il web utilizzato come modalità di raccolta dei dati comporta alcuni rischi: i consumatori avvertono un senso di sicurezza e non sono solitamente a conoscenza dell'enorme quantità di informazioni che rendono pubbliche. Questi dati comprendono cronologia di navigazione, informazioni estratte dai social media e cronologia degli acquisti, solo per citarne alcuni. Attraverso identificatori quali indirizzo mail o indirizzo IP, il consumatore viene identificato univocamente e ciò consente di creare un profilo estremamente dettagliato e ricco di informazioni. [43] Inoltre, grazie a questi codici univoci, a seguito di una richiesta di eliminazione o di anonimizzazione, i dati possono essere ricollegati nuovamente al consumatore, rendendo vana la procedura. [44] La profilazione può diventare molto invasiva, e il consumatore è spesso inconsapevole di questa problematica.



**Politecnico
di Torino**

Le implicazioni sono molteplici. La scarsa accuratezza dei dati potrebbe causare danni alla reputazione, [43] e l'erosione della fiducia sociale: l'utente, a causa della mancanza di trasparenza e di controllo sui propri dati, potrebbe perdere fiducia nelle istituzioni. Inoltre, la disponibilità e l'accumulo di enormi quantità di informazioni personali comporta un incremento del rischio di violazioni della sicurezza e di furto di identità, anche ma non solo a seguito di infiltrazioni all'interno dei database delle aziende. [45]

L'utilizzo di dati personali come merce di scambio comporta lo sviluppo di un'economia della sorveglianza, con implicazioni non facilmente prevedibili. [46]

La raccolta delle informazioni non si ferma a ciò che viene reso pubblico online: le fonti dei data broker sono varie e spesso le aziende tentano di non dichiararle esplicitamente: ciò rende complicato capire come e dove le informazioni vengono raccolte.

2.2 Competition

Nel contesto economico, quando si parla di competition si intende la rivalità tra le diverse imprese operanti in un determinato settore. Questa rivalità sussiste al fine di appropriarsi di diverse tipologie di vantaggi economici, quali ad esempio quote di mercato e profitti.



**Politecnico
di Torino**

La competizione comporta che le aziende siano incentivate a ridurre i prezzi per attirare nuovi clienti. Come si vede dal grafico, all'equilibrio il surplus del consumatore è all'ottimo, e quindi la differenza tra il prezzo massimo che offrirebbe e quello effettivamente pagato è più alta. La condizione di equilibrio è raggiunta in questo caso quando la competition è perfetta e quando il costo marginale per ottenere un'altra unità dal fornitore è uguale al prezzo più alto al quale un cliente comprerebbe un'unità. [47]

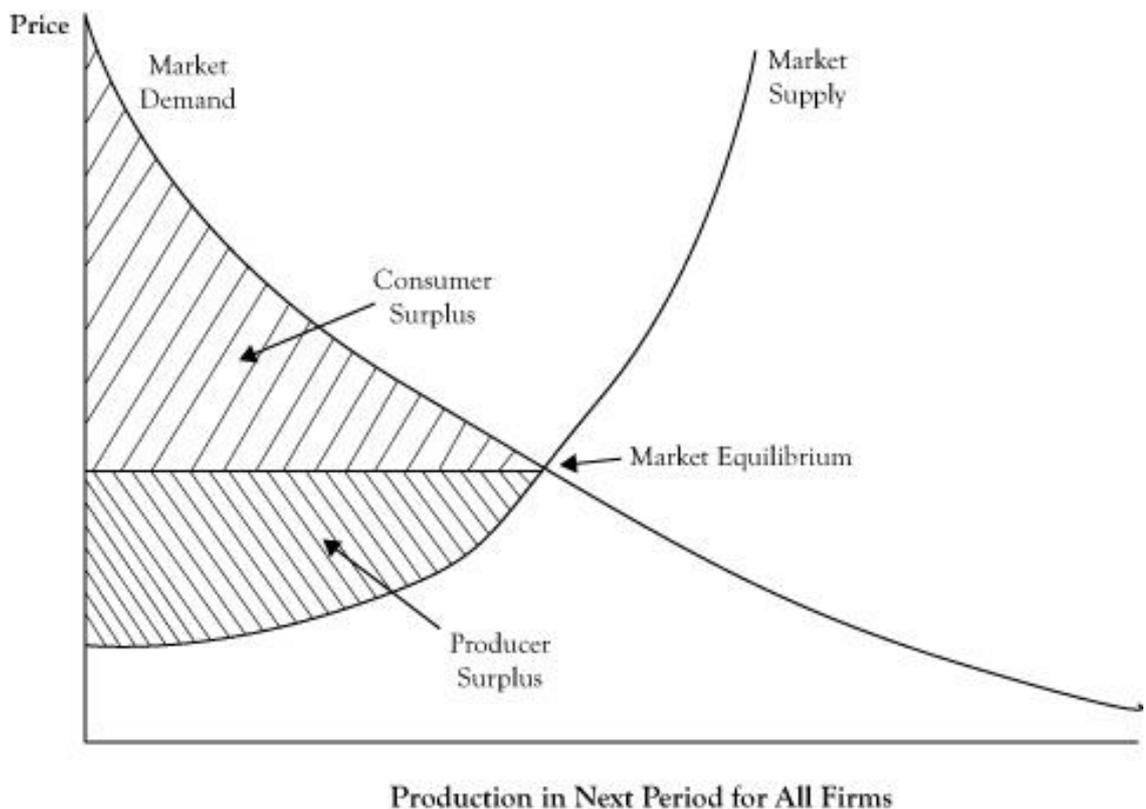


Figura 5: Grafico di suddivisione surplus all'equilibrio [47]



**Politecnico
di Torino**

Al contrario, in una situazione di competitività non perfetta, l'azienda può permettersi di alzare il prezzo oltre al valore del prezzo di equilibrio. Ciò comporta un incremento del surplus dell'azienda, a discapito di una riduzione del surplus del consumatore. Questo scambio non è però proporzionato: si crea quindi una perdita netta, non risolvibile nemmeno tramite l'imposizione di tasse al fine di redistribuire il surplus. Da ciò si evince il fatto che il punto di equilibrio non sia solamente il punto di incontro tra domanda e offerta, ma anche il punto nel quale il surplus totale è massimo.

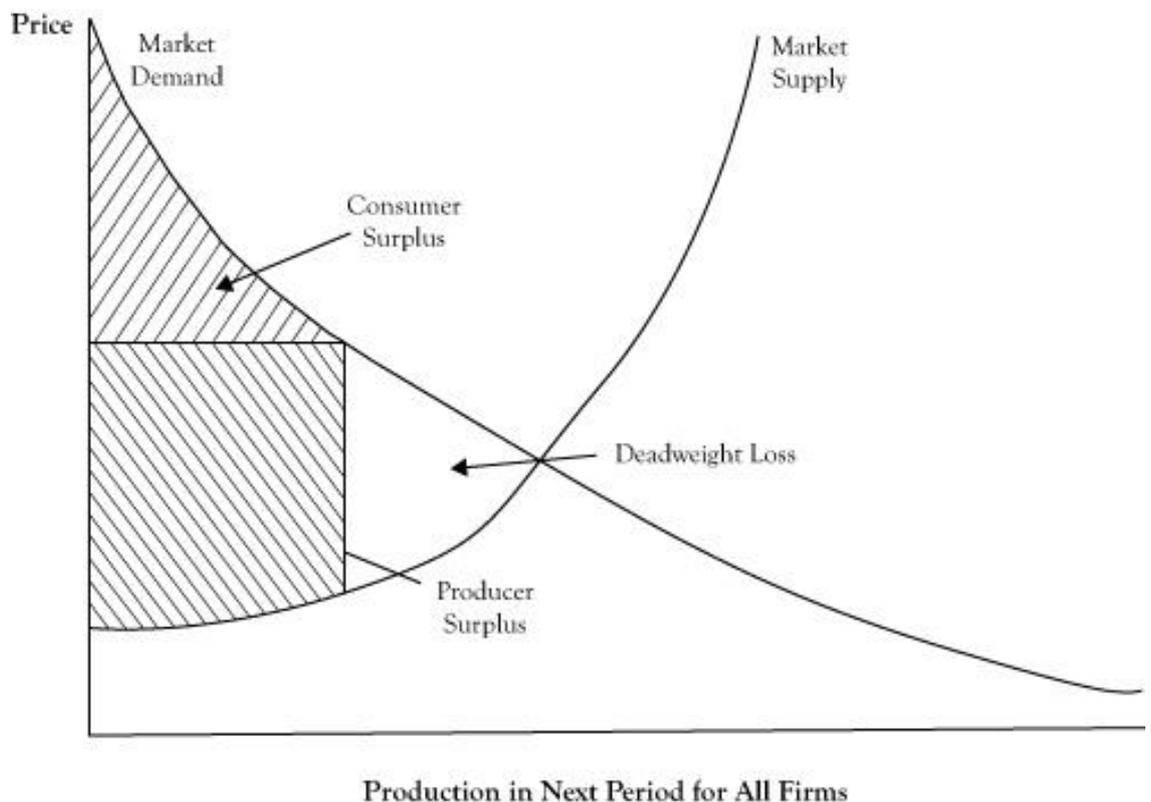


Figura 6: Grafico di suddivisione surplus non all'equilibrio [47]



**Politecnico
di Torino**

Nel contesto del mercato dei data broker, ciò avviene con diverse conseguenze. Nel caso del fallimento di mercato (condizione che si verifica quando si ottiene una perdita netta) il surplus del consumatore è inferiore quando il data broker partecipa al mercato. Infatti, questo compenserebbe una ricompensa minima al consumatore, il quale riceverebbe inoltre un servizio peggiore da parte dei fornitori di contenuti e servizi online (social media, piattaforme di streaming, siti di e-commerce).

Al contrario, nel caso in cui il data broker sia efficiente e non si verifichi un fallimento di mercato, i consumatori solitamente beneficiano della sua presenza. [48]

A prova del bilancio tra competition e surplus del consumatore, uno studio effettuato in Finlandia evidenzia che, a seguito delle cessazioni delle operazioni di un data broker locale, i prezzi nel settore della vendita al dettaglio di prodotti alimentari siano aumentati nel breve termine, a causa della diminuzione di competitività del settore. [49]

2.3 Errori nel profiling

Come detto precedentemente, i data broker ottengono le informazioni relativi ai consumatori tramite molteplici canali e fonti. Utilizzano queste informazioni



**Politecnico
di Torino**

per la creazione di profili, ma non sempre i dati ottenuti sono corretti: il 40% degli attributi riguardanti i consumatori sono imprecisi o errati. [50] I data broker più virtuosi in questo campo arrivano al 50% di record corretti. [51]

Ciò comporta molteplici rischi, tra i quali la discriminazione di alcune categorie ed errori nel calcolo del credit score: si tratta di un punteggio utile a misurare l'affidabilità creditizia di un consumatore, ovvero la probabilità che questo rimborsi puntualmente il prestito. Questo punteggio è calcolato sulla base di diverse informazioni, ed è utilizzato da istituzioni finanziarie, quali banche e assicurazioni.

Un'analisi della Federal Trade Commission nel 2013 ha evidenziato come il 26% dei dati creditizi dei consumatori fossero errati e avrebbero potenzialmente causato un aumento delle spese creditizie per gli utenti. [7] La correzione di questi errori ha portato al cambiamento del credit score nel 10% dei casi. [52]

La frequenza di valori errati è probabilmente dovuta alla modalità di calcolo del punteggio: secondo uno studio tedesco, infatti, il valore del credit score è spesso una stima, e la sua valenza sul piano individuale dovrebbe essere ridimensionata. Un esempio di una modalità di approssimazione del credit score è quella utilizzata dall'azienda Cignifi, la quale per calcolarlo utilizzava dati quali la durata delle chiamate, chi le effettuava e chi le riceveva. [7] Alcuni episodi di scarsa affidabilità del credit score possono essere attribuiti ad errori di identificazione: può accadere che ad un consumatore venga attribuito il credit score di un altro. [21]



**Politecnico
di Torino**

Nonostante la comprovata scarsa precisione di questo indicatore, esso viene comunque utilizzato da moltissime istituzioni finanziarie, e la mancanza di un credit score o il fatto di avere un punteggio non ottimo può precludere l'accesso di alcuni consumatori ad alcuni servizi, e anche escluderli dal sistema finanziario. [21]

La correzione degli errori non è però semplice: i consumatori si devono rivolgere ai data broker, chiedendo di visionare l'informazione, e, se incorretta, possono chiederne la revisione. Ciò però non comporta l'immediata risoluzione del problema: possono volerci anni per vedere la fine del processo. [50]

2.4 Discriminazione di categorie a rischio

La discriminazione di persone a rischio è un problema rilevante: questa si verifica quando alcuni individui subiscono un trattamento svantaggioso o comunque ineguale a causa della loro appartenenza ad alcune categorie sociali. Queste possono essere di differenti tipologie, quali ad esempio categorie etniche, religiose, di genere o di orientamento sessuale. Le conseguenze possono essere molteplici: le opportunità possono essere razionate, i prezzi discriminati (ad esempio sulla base di dati sulla salute alcune categorie potrebbero pagare premi di assicurazione più elevati [53]). Possono essere inoltre esacerbate problematiche quali la marginalizzazione e l'esclusione sociale.



**Politecnico
di Torino**

È possibile suddividere la problematica in due differenti tipologie di discriminazione: quella diretta e quella indiretta. Nel primo caso la discriminazione è basata su attributi sensibili, e può concretizzarsi in diversi modi: ad esempio, la profilazione etnica può incidere sull'accesso a soluzioni di affitto di appartamenti e sull'utilizzo arbitrario di sorveglianza. La discriminazione indiretta invece avviene a causa delle derive di procedure apparentemente neutre. Queste possono talvolta discriminare involontariamente un gruppo: ad esempio, il codice postale di una determinata zona può fungere da alternativa ad un attributo di tipologia razziale, e portare un dato apparentemente non discriminatorio a causare discriminazione abitativa. [54]

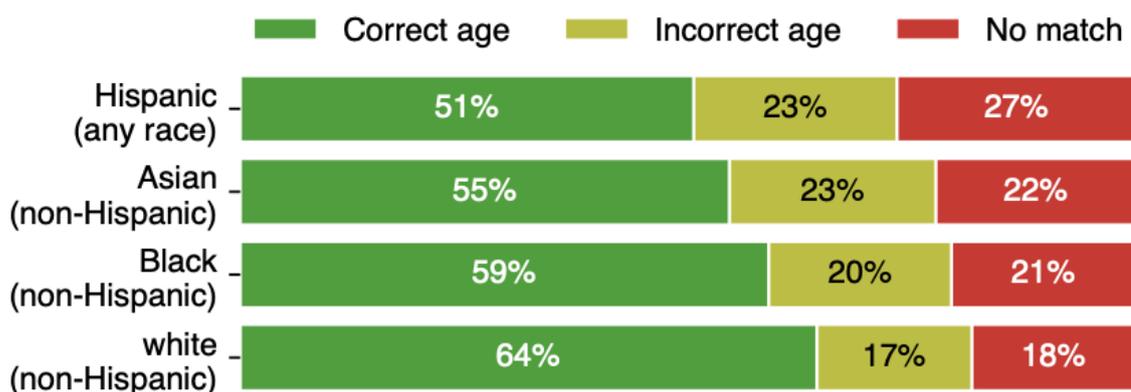


Figura 7: Grafico di accuratezza dei dati in base a discriminante etnica [50]



**Politecnico
di Torino**

Quello evidente nel grafico è il problema della sottorappresentazione: la popolazione di bianchi non ispanici ha una correttezza dei dati maggiore rispetto alle altre: nel grafico l'esempio è quello dell'età. La sottorappresentazione, oltre all'incorrettezza dei dati, può anche portare ad una discriminazione nell'utilizzo dei dati biometrici: la percentuale di errori nel suo utilizzo è più alta nella popolazione degli americani di origine africana. [7]

L'evidenziazione dei rischi e della loro materializzazione concreta è evidente guardando la frequenza delle notizie di scandali, tra i quali si possono considerare i data breach, ovvero infiltrazioni di persone all'interno dei database dei data broker, al fine di sottrarre informazioni di milioni di consumatori.

- Equifax: nel 2017 Equifax è stata vittima di una delle più grandi violazioni di dati nella storia USA: circa 145 milioni di consumatori sono stati interessati da questo furto. Per circa un mese il database è stato soggetto a intrusioni: l'azienda ha rivelato di esserne stata vittima dopo diverso tempo. A seguito di questi avvenimenti, sono state intentate diverse class action ai danni dell'azienda. [55]
- Exactis: nel 2018 sono state esposte senza adeguata protezione 340 milioni di informazioni personali: l'azienda Exactis le ha infatti memorizzate senza alcun firewall, ed erano accessibili tramite ricerche mirate. Exactis, dopo aver appreso dell'esposizione del database e dopo averlo adeguatamente protetto, non ha mai rivelato se i dati siano stati trafugati. [56]



**Politecnico
di Torino**

- Verifications.io: nel 2019 circa 763 milioni di record personali sono stati trafugati dal servizio di verifica e-mail. Ciò è avvenuto a causa di un database lasciato non protetto da alcun sistema di autenticazione, all'interno dell'applicativo MongoDB. [57]
- SafeGraph: nel 2020 questa azienda che opera nel mercato dei dati di geolocalizzazione è stata coinvolta in controversie, a causa della commercializzazione di dati che riguardano le visite a cliniche per l'aborto. L'azienda ne entrava in possesso grazie ad app comuni installate negli smartphone dei consumatori, e venivano cedute ad aziende terze. [58]
- National Public Data: nel 2024, l'azienda NPD, fornitrice di servizi di prevenzione di frodi e di controllo dei precedenti, ha subito un furto potenziale di 2,9 miliardi di informazioni sensibili riguardanti cittadini americani, inglesi e canadesi. I dati sono stati divulgati sul dark web, e comprendevano nomi, numeri di previdenza sociale e indirizzi. [59]

3 Benefici

Le implicazioni dei data broker non sono solamente negative: essi apportano, oltre ai già citati rischi, alcuni benefici ai consumatori.



**Politecnico
di Torino**

3.1 Targhettizzazione

Uno degli aspetti positivi del mercato dei dati è il miglioramento della targhettizzazione: grazie alla creazione di profili ricchi di informazioni sui consumatori, e successivamente alla clusterizzazione di questi profili, gli utenti possono ricevere pubblicità personalizzate sui loro interessi e sulle loro caratteristiche. Ciò comporta alcuni lati positivi.

- **Rilevanza e personalizzazione:** le aziende possono inviare messaggi pubblicitari più pertinenti ad ogni tipologia di consumatore: servizi locali grazie ai dati sulla geolocalizzazione, ma anche prodotti pertinenti al loro stile di vita e potenzialmente alla loro capacità di spesa, basandosi sulle informazioni creditizie. Inoltre, gli interessi e i valori personali da essi estraibili possono permettere l'offerta o l'esclusione di alcune tipologie di prodotti.
- **Riduzione del rumore pubblicitario:** la quantità di pubblicità irrilevante e non interessante per il consumatore diminuisce, e così migliora l'esperienza utente dei servizi online.
- **Offerte pertinenti:** grazie ai dati sulla navigazione, i data broker possono proporre al consumatore offerte riguardanti prodotti ai quali questo aveva in precedenza mostrato interesse, permettendogli di ottenere un risparmio.



**Politecnico
di Torino**

3.2 Prevenzione delle frodi

La prevenzione delle frodi è uno degli aspetti positivi generati dai data broker, e a beneficiarne possono essere cittadini privati ma anche aziende o enti statali. Grazie all'enorme mole di dati in loro possesso, i broker di dati possono rilevare incongruenze nelle informazioni personali durante svariate tipologie di processi. Ad esempio, nel corso di un tentativo di apertura di un contratto telefonico, se il codice fiscale o il Social Security Number non combacia con il proprietario della carta di credito, il sistema è in grado di segnalare la transazione. Una procedura simile può essere utilizzata per verificare la corrispondenza tra i dati di una carta di credito con quelli di spedizione, ma anche per controllare eventuali discordanze tra il reddito dichiarato e gli acquisti effettuati. [1]

Alcuni servizi forniti dagli stessi data broker sono: Experian IdentityWorks, Equifax ID Watchdog e TransUnion TrueIdentity [60]. Questi operano controllando e raccogliendo dati da database, agenzie di credito e dark web, analizzandoli e associandoli ai rispettivi consumatori. Nel caso di attività sospette, inviano avvisi e lo notificano al cliente, consentendogli di bloccare l'operazione.

Le frodi si estendono a qualsiasi contesto, e ad esempio in quello della sanità americana, dal 3 al 10% del suo costo totale è attribuibile a casi di frode. Ciò può essere mitigato dall'utilizzo di grandi dataset per controllare la presenza di anomalie o schemi ricorrenti: la creazione e l'utilizzo di liste di esclusione, ovvero



**Politecnico
di Torino**

medici o operatori sanitari segnalati in precedenza, le cui operazioni vengono poi segnalate. [61]

3.3 Verifica dell'identità

La verifica dell'identità è un servizio utile per consentire ai consumatori di accedere e velocizzare diverse procedure. Un esempio può essere quello dell'apertura di un conto corrente, o la richiesta di un servizio finanziario, quale prestito, mutuo o assicurazione tramite canali online. I dati inseriti dal consumatore vengono infatti verificati e confrontati con quelli presenti nel database: uno dei tanti servizi a fornire questo tipo di controlli è "Precise ID Platform", del data broker Experian. Grazie infatti alla grande quantità di dati in loro possesso, la procedura di autenticazione dell'utente è facilitata: l'accesso da una località nota o da un dispositivo riconosciuto facilitano la procedura.

Fondamentale ormai al fine della verifica dell'identità è l'uso di una caratteristica personale o fisica univoca. È il caso dei riconoscimenti biometrici, usati ormai in molti ambiti, governativi e commerciali.



**Politecnico
di Torino**

4 Stato attuale delle policy

Per poter garantire il corretto funzionamento dell'economia, la regolamentazione dei mercati è fondamentale, e quello dei data broker non fa eccezione: trattandosi di una tipologia di attività che coinvolge i consumatori, si è rivelata importante anche la tutela della loro privacy. Differenti stati hanno però diverse tipologie di approcci, a seconda di quanto la società e la politica che la rappresenta ritengano che l'intervento dello stato debba essere radicale. Comprendere quindi le differenti normative adottate dai regolatori è fondamentale per capire il contesto nel quale si muovono i data brokers. Questo settore è infatti pesantemente influenzato dalle regolamentazioni, le quali stabiliscono le operazioni lecite e quelle non lecite.

I due regolatori presi in esame sono gli Stati Uniti d'America e l'Europa. Nel primo caso, l'approccio normativo è caratterizzato dalla mancanza di una regolamentazione centralizzata a livello federale: in sua assenza, alcuni stati hanno legiferato in merito, ma ciò ha contribuito a creare un panorama almeno parzialmente frammentato. [21] Per quanto riguarda invece l'Unione Europea, questa ha adottato un sistema di regolamentazione più centralizzato.

4.1 Situazione in Europa



**Politecnico
di Torino**

L'Unione Europea si fonda su un approccio orientato ai diritti e alla protezione della privacy dei cittadini. Questa politica viene espressa nel Regolamento Generale sulla Protezione dei Dati, noto con la sigla GDPR. Questo regolamento si pone come obiettivo la creazione di un mercato unico dei dati fondato sulla trasparenza tramite l'imposizione di limiti precisi alle azioni dei data broker, per conferire parte del controllo dei propri dati ai cittadini. [45] Le regolamentazioni del GDPR possono essere divise in due sezioni principali: quelle riguardanti la privacy e quelle sul trattamento dei dati.

Uno dei vincoli imposti al mercato dei dati è l'obbligo del consenso esplicito e informato: i data broker sono tenuti a fornire informazioni sulla finalità della raccolta dei dati e sui diritti dei soggetti interessati, e il consenso implicito o presunto non è più consentito. Questo requisito comporta l'impossibilità da parte del data broker di raccogliere dati senza che i consumatori ne siano consapevoli. [53]

Viene inoltre regolamentato lo scopo di raccolta dei dati: ogni data broker deve infatti dichiararne la finalità. Questa deve essere specifica, legittima ed esplicita: ciò pone un freno alla raccolta indiscriminata di ogni informazione disponibile e anche al loro utilizzo per finalità diverse da quelle stabilite: nel caso in cui un data broker voglia sfruttare un dato per un altro fine, dovrà raccogliere un nuovo consenso. [46]

Uno dei diritti acquisiti dai consumatori grazie al GDPR è il diritto al controllo sui propri dati: ogni soggetto interessato può richiedere la cancellazione dei propri



**Politecnico
di Torino**

dati, così come può richiederne la correzione in caso di errori e la portabilità: può quindi trasferire l'informazione ad un altro data broker. [62] Nella cancellazione, però, non è imposta l'interruzione della raccolta: il data broker potrebbe quindi riprenderla dopo un periodo di tempo. [53]

Il trattamento dei dati è un altro dei campi regolamentati dal GDPR: i data broker sono infatti tenuti a mantenere alti livelli di trasparenza: essi devono poter dimostrare la fonte, la modalità e la finalità dei dati raccolti, quando interpellati dalle autorità di controllo. Le informazioni non possono essere trasferite a paesi terzi che non offrono un adeguato standard di protezione. Viene inoltre imposto un elevato livello di sicurezza: il titolare del trattamento, ovvero il data broker, è considerato responsabile del rispetto delle regole della sicurezza, per evitare perdite, modifiche o accessi non autorizzati. Il data broker deve quindi dimostrare di poter fornire garanzie sufficientemente elevate. [45]

Le violazioni delle norme contenute all'interno del GDPR comportano sanzioni rilevanti, che possono arrivare sino al valore più elevato tra il 4% del fatturato annuo dell'azienda e 20 milioni di euro. [45]

Il GDPR non è l'unico strumento dell'Unione Europea per la regolamentazione del mercato dei dati.

- Data Governance Act (DGA): stabilisce limiti all'utilizzo e alla condivisione delle informazioni. [63] Esso non si focalizza specificatamente sui data broker: i dati interessati sono quelli non personali, ma dato che spesso questi vengono aggregati a quelli personali, ha un impatto significativo sulle attività



**Politecnico
di Torino**

dei broker di dati. La condivisione delle informazioni tra settori diversi viene facilitata, così come l'accessibilità dei dati del settore pubblico, il cui utilizzo si ritiene possa apportare benefici. All'interno del documento viene incoraggiato l'altruismo dei dati: le organizzazioni possono iscriversi come organizzazioni per l'altruismo dei dati riconosciute nell'UE, inquadramento per le società che desiderano raccogliere dati a fini di interesse pubblico. [62]. Il DGA prevede inoltre la designazione di un'autorità competente che possa supervisionare l'applicazione del regolamento. Essa deve essere indipendente dalle aziende fornitrici di servizi per la condivisione dei dati, ma anche dalle organizzazioni per l'altruismo dei dati. [5]

- Digital Services Act (DSA): questa regolamentazione impone obblighi di due diligence, ovvero l'imposizione di misure al fine di garantire sicurezza e affidabilità dei servizi forniti dai data broker. Essi devono inoltre fornire trasparenza riguardo le proprie politiche di policy interne e la designazione di rappresentanti legali, nel caso in cui si tratti di data broker con sede esterna all'Unione Europea. [64] [65]
- Digital Markets Act (DMA): questa regolamentazione è mirata ai gatekeeper, ovvero grandi piattaforme che minano l'equità dei mercati digitali abusando della loro posizione dominante nel mercato. Questi sono infatti tenuti ad ottenere consenso esplicito per il trattamento dei dati personali, ma anche a limitare l'utilizzo di essi, nel caso in cui i data broker siano fornitori e di servizi ad aziende ma anche loro competitor. È questo il caso ad esempio di



**Politecnico
di Torino**

Amazon, compagnia fornitrice di servizi per aziende di vendita al dettaglio, ma al contempo è considerabile retailer diretta. [66] [64]

4.2 Situazione negli USA

A differenza dell'approccio dell'Unione Europea, la regolamentazione del mercato dei data broker negli Stati Uniti non è unificata. Si presenta infatti come un mosaico di leggi settoriali, e ciò comporta un elevato livello di complessità e una tutela dei consumatori non completa. Non esiste infatti una legge federale omnicomprensiva sulle azioni dei data broker, ma vi sono numerose normative focalizzate su alcune tipologie di dati. [21]

- ◆ FCRA: la Fair Credit Reporting Act si impone come obiettivo la regolamentazione di raccolta, divulgazione e uso dei dati creditizi dei consumatori. Viene applicata in particolare alle agenzie di credito, e stabilisce che i dati debbano rispettare standard di accuratezza e che possano essere controllati e corretti dai consumatori. [1]
- ◆ GLBA: la legge Gramm-Leach-Bliley Act (GLBA) regola le istituzioni finanziarie e la protezione dei dati dei propri clienti: questi ultimi devono infatti essere avvisati delle politiche sulla privacy della società, e devono poter fornire il proprio consenso alla condivisione dei dati con enti terzi. [67]



**Politecnico
di Torino**

- ◆ COPPA: In materia di dati di bambini inferiori ai tredici anni è la legge COPPA (Children's Online Privacy Protection Act) ad intervenire: ogni sito web o servizio online deve ottenere il consenso dei genitori per poter raccogliere o divulgare i dati dei bambini. [21]
- ◆ FTC Act: Un organo rilevante per la regolamentazione del trattamento dei dati è la Federal Trade Commission, FTC, la quale si pone come obiettivo la protezione dei consumatori e la promozione della concorrenza. Si tratta di un organo che conduce numerosi studi e indagini di settore, per poi avviare azioni legali contro le aziende che non rispettano la FTC Act. Tramite questa legge vengono infatti vietate pratiche commerciali sleali, ma non viene fornito un quadro normativo completo. [68]

Per sopperire alla mancanza di legislazioni federali specifiche e omnicomprendenti sul mercato dei data broker, alcuni stati si sono mossi indipendentemente, promulgando leggi come il California Consumer Privacy Act (CCPA) e una sua successiva modifica, il California Privacy Rights Act (CPRA). [53] Queste due leggi tutelano sotto diversi aspetti i diritti dei consumatori. Tra le altre cose, infatti, è garantita la possibilità di conoscere quali dati siano stati raccolti, così come l'eventuale cancellazione su richiesta. Inoltre, le aziende possono vedersi negata la possibilità di vendere e utilizzare queste informazioni, e sono tenute a correggerle, se il consumatore lo richiede. Lo sfruttamento dei dati sensibili può essere ulteriormente limitato da una richiesta del soggetto interessato. [69] [70]



**Politecnico
di Torino**

Al contrario delle normative federali, le leggi promulgate dallo stato della California interessano la totalità del mercato dei data broker, e non si limitano a regolamentare una tipologia specifica di azienda o di dati.

Vi sono stati e vi sono tutt'ora tentativi di legiferare in materia di data broker: nel 2022 è stata formulata la American Data Privacy and Protection Act, ADPPA. Si tratta di una proposta di legge che si pone come obiettivo la creazione di uno standard nazionale per la privacy dei dati. La legge non è però stata approvata dal congresso. [71]

Nel 2024, invece, ha visto la luce l'APRA (American Privacy Rights Act). Le sue finalità sono simili a quelle della ADPPA, ma con alcune differenze nella formulazione: l'APRA considera dati personali una varietà maggiore di informazioni, quali ad esempio dati biometrici, di geolocalizzazione e di attività online. Ciò consentirebbe la regolamentazione dei diritti dei cittadini e degli obblighi delle aziende anche su queste tipologie di informazioni. Inoltre, l'APRA consentirebbe l'applicazione delle leggi statali in materia, e un controllo maggiore: imporrebbe consensi obbligatori da parte del consumatore, e le regolamentazioni verrebbero applicate a tutti i data broker, senza le esenzioni presenti nell'ADPPA. [71] [72]

4.3 Problematiche delle regolamentazioni



**Politecnico
di Torino**

Nonostante le diversità nell'approccio e nel risultato, una caratteristica comune tra la legislazione Europea e quella Americana sono le limitazioni applicative. Infatti, sia nel GDPR che nelle proposte di leggi federali, si dà grande importanza al consenso informato e alle azioni intraprese dai consumatori. Ciò è però reso poco efficace dalla elevata difficoltà ad esercitare il controllo sulle proprie informazioni: tipicamente il consumatore è scarsamente informato su esistenza e attività dei data broker, e non è quindi in grado di esercitare i propri diritti. Il consenso informato rappresenta una problematica anche per i data broker: il loro modello di business si basa spesso sulla raccolta di dati e solo successivamente sulla loro analisi, e ciò rende difficile per le aziende conoscere la finalità dei dati raccolti prima ancora di ottenerli. [43]

La rimozione dei dati deve essere richiesta da parte del singolo consumatore verso il singolo data broker, e ciò rende inefficiente ed improbabile la richiesta. Inoltre, una volta eliminati i dati questi possono essere raccolti nuovamente dopo un certo periodo di tempo. [53]

La maggiore forza legislativa dell'Unione Europea in materia e la politica di rifiuto di operazioni di fusioni tra grandi aziende, volta a evitare la creazione di oligopoli, potrebbe porre ostacoli alla creazione di aziende forti e capaci di competere a livello globale. [73]

USA ed Europa non sono le uniche entità che si sono trovate a dover controllare le attività dei data broker e a doverne gestire le problematiche: anche l'Australia, infatti,



**Politecnico
di Torino**

tramite la Commissione australiana per la concorrenza e i consumatori (ACCC), all'interno di un'indagine sulle piattaforme digitali, si è soffermata sulle pratiche commerciali messe in atto dai data broker. Infatti, sono state esplorate le modalità tramite le quali i data broker raccolgono e sfruttano i dati, eventuali problematiche legate a concorrenza e tutela del consumatore, e in particolare, ha espresso preoccupazioni concernenti la trasparenza. [74]

La finalità di tale indagine non è quella di risoluzione del problema, ma essa si pone come base per ulteriori indagini e legislazioni in merito.

In un documento posto all'attenzione all'ACCC nel 2024, il gruppo di tutela dei consumatori CHOICE ha messo evidenziato vuoti normativi nell'ambito della privacy, includendo suggerimenti per la loro risoluzione. Tra questi, l'allargamento dei confini delle informazioni definite personali, l'aggiunta di requisiti di ottenimento del consenso prima di poter svolgere determinate attività e obbligo di ogni data broker di dichiarare la finalità della raccolta dei dati. [75] Tale approccio, se implementato, potrebbe essere più simile a quello europeo che a quello americano.

Un ulteriore esempio di legislazione sui data broker e la privacy è la PIPL, Personal Information Protection Law, promulgata dalla Repubblica Popolare Cinese nel 2021. Questa, infatti, si applica alle organizzazioni che raccolgono e sfruttano informazioni personali identificabili (PII) in Cina e all'estero. Richiede che i consumatori possano correggere e in generale prendere decisioni riguardo l'utilizzo dei propri PII. Pone, inoltre, vincoli sulla commercializzazione all'estero di tali dati, ingiungendo restrizioni aggiuntive e limitazioni del loro utilizzo. [76] [77] Si tratta di un approccio



**Politecnico
di Torino**

paragonabile a quello europeo, pur vigendo in un contesto politico e di tutela dei cittadini totalmente differente.

Conclusioni

La rilevanza delle aziende operanti sotto la definizione di data broker è evidente: un intero settore economico si fonda sul commercio dei dati dei consumatori, per la maggior parte ignari di ciò. Lo scopo di questa tesi è quello di analizzare il rapporto rischi-benefici per il cittadino, e di descrivere alcune delle regolamentazioni più rilevanti in materia di commercio dei dati. È stata fornita una definizione delle aziende di data brokering, per poi entrare nello specifico per quanto riguarda il loro modello di business: si è quindi trattato delle fonti dei dati. Queste sono varie, e spaziano dai documenti pubblici ai social media, dalle più generiche attività online alle abitudini commerciali dei consumatori: molte delle azioni compiute da essi vengono interessate dai data broker, che ne entrano in possesso.

In seguito è stata descritta la modalità attraverso la quale i dati vengono trattati. Essi, infatti, possono essere memorizzati grazie a tecnologie differenti tra loro per quanto riguarda sicurezza e economicità. Sono state considerate alcune tipologie di dati, i quali attraversano diversi passaggi, sino ad arrivare al clustering dei profili creati. La trattazione ha preso in esame le due modalità principali di cessione dei dati, ovvero la vendita e la concessione in licenza, utilizzate dalle diverse aziende a seconda delle finalità dei dati e del business model del data broker. Successivamente è stato preso



**Politecnico
di Torino**

in esame il settore in questione e le sue caratteristiche: la crescita passata è stata elevata, pari al 6,5% annuo dal 2012 al 2019, e per il prossimo periodo si prevede un'ulteriore espansione dell'8% annuo. Gli Stati Uniti d'America sono l'area geografica più rilevante nel mercato dei data broker, seguita dall'Europa e da alcuni mercati emergenti, quali ad esempio Cina e India.

Tramite l'analisi delle cinque forze di Porter si è giunti ad alcune conclusioni. Le barriere all'entrata relativamente alte e la concentrazione del mercato comportano il fatto che la competizione temuta dalle aziende è quella interna al mercato, più che al possibile arrivo di nuovi competitor. Si è inoltre stabilito che si tratta di un settore complesso e competitivo, con economie di scala rilevanti.

Nonostante la loro indubbia rilevanza e nonostante posseggano dati su potenzialmente la totalità dei cittadini, le entità operanti in questo settore sono totalmente sconosciute al grande pubblico. Sono state quindi elencate e brevemente analizzate le aziende più rilevanti a livello globale e la loro storia.

Nel secondo capitolo sono stati esaminati i rischi che le operazioni dei data broker pongono sui consumatori. Questi sono la raccolta dei dati online, la quale comporta l'erosione della privacy, ma anche la situazione di competition non perfetta, con un conseguente abbassamento del livello del surplus del consumatore. Anche gli errori di profiling, che caratterizzano il 40% dei dati totali, possono causare gravi problematiche. In particolare, il dato del credit score è errato nel 26% dei casi. Ciò può comportare il mancato accesso a servizi ai quali il consumatore avrebbe diritto o un aumento del costo di prestiti o mutui. La discriminazione delle categorie a rischio è un problema messo in evidenza, tra gli altri, da uno studio che evidenzia il grado di



**Politecnico
di Torino**

accuratezza dei dati a seconda della discriminante etnica: in questo caso, alcune etnie sono più penalizzate di altre. Una delle modalità di concretizzazione dei rischi è la diffusione pubblica dei dati, e ciò viene preso in considerazione tramite 5 tra i maggiori data leaks avvenuti negli ultimi 6 anni.

Nel terzo capitolo vengono analizzati i benefici apportati ai consumatori grazie ai data broker. La targhettizzazione è più precisa, e ciò riduce la pubblicità irrilevante e migliora l'esperienza utente. Inoltre, grazie ai servizi messi a disposizione dagli stessi data broker, il consumatore beneficia della riduzione della probabilità di essere vittima di frodi e di furti di identità.

Il bilanciamento tra rischi e benefici può essere considerato sfavorevole: le conseguenze dei rischi possono avere un impatto considerevole, e non sempre sono note al consumatore. Il più evidente è senza dubbio il rischio dell'erosione della privacy, ma probabilmente non si tratta di quello con conseguenze più gravi. La percentuale di errori nel dato del credit score, infatti, evidenzia una situazione critica: le istituzioni finanziarie si basano su questo dato per concedere o meno un prestito a un consumatore, e lo usano anche per determinare il costo del debito dello stesso. Ciò comporta un enorme rischio a carico del cliente nel caso di un credit score sottostimato, ma anche un rischio a spese dell'ente finanziatore, nel caso di un valore erroneamente alto. Non trascurabile è anche il rischio di discriminazione di categorie a rischio: l'attività dei data broker può avere come conseguenza il mancato rispetto del diritto abitativo di un consumatore sulla base ad esempio di caratteristiche etniche.



**Politecnico
di Torino**

D'altra parte, i benefici, in particolare la riduzione di rischi di frodi e furti di identità, non sono ottenibili esclusivamente tramite l'utilizzo dei servizi forniti dai data broker, e le soluzioni alternative sono spesso funzionanti e adottate dalla maggior parte dei consumatori attenti a queste tematiche. In aggiunta, tali benefici possono essere considerati delle soluzioni per mitigare rischi generati, almeno parzialmente, dalla stessa raccolta e gestione dei dati. Spesso, infatti, frodi e furti di identità vengono perpetrati grazie a dati e informazioni che sono divenute accessibili ad attori malintenzionati.

Le operazioni dei data broker sono varie e il rischio che diventino opache è concreto: ciò comporta la necessità che i regolatori intervengano per gestirne le problematiche. Questo è l'argomento trattato nel quarto capitolo, nel quale le diverse legislazioni vengono esaminate: in primis la situazione in Europa, con GDPR, DGA, DSA, DMA e in seguito le leggi negli USA.

È lecito stabilire che l'Unione Europea ha adottato un quadro normativo più stringente: negli USA a livello federale le leggi riguardanti i data broker sono quasi assenti, e quelle più omnicomprensive devono ancora essere discusse e approvate dal congresso. Le regolamentazioni specifiche vigono a livello dei singoli stati, e sono quindi difficilmente applicabili e limitate. Alla situazione attuale, all'interno del mercato USA, le limitazioni alle attività dei data broker sono quasi assenti. Ciò consente l'attuazione di pratiche, tra le quali la raccolta e la vendita di dati senza consenso, a qualunque ente li richieda, sia esso nazionale o estero.

Per quanto riguarda la situazione Europea, invece, pur avendo agito implementando regolamentazioni comunitarie, queste non garantiscono un completo e totale diritto



**Politecnico
di Torino**

al controllo sui propri dati. Inoltre, le differenti legislazioni vigenti al livello dei singoli stati potrebbero fornire scappatoie alle regolamentazioni Europee, indebolendole ulteriormente.

D'altro canto, alcune delle imposizioni sui data broker potrebbero essere difficili da applicare: ne è un esempio l'obbligo di dichiarare la finalità dei dati raccolti, informazione spesso non in possesso del data broker. Questo potrebbe fungere da deterrente e limite per innovazione e investimenti nel settore: è quindi evidente l'importanza di trovare un equilibrio funzionale tra disciplina legale e attrattività del mercato.



**Politecnico
di Torino**

Bibliografia

- [1] "U.S. FTC," 2014. [Online]. Available: U.S. Federal Trade Commission, Data brokers: a call for transparency and accountability.
- [2] European Data Protection Board, "Linee guida 07/2020 sui concetti di titolare del trattamento e di responsabile del trattamento ai sensi del GDPR," 2020. [Online]. Available: https://www.edpb.europa.eu/system/files/2023-10/edpb_guidelines_202007_controllerprocessor_final_it.pdf.
- [3] "U.S. Senate Committee on Commerce, Science, and Transportation, A review of the data broker industry: collection, use, and sale of consumer data for marketing purposes," 2012.
- [4] Congresso USA, "Freedom of Information Act".
- [5] Norton, "Data Governance Act," [Online]. Available: <https://www.nortonrosefulbright.com/en-gb/knowledge/publications/8102ac08/data-governance-act>.
- [6] I. Lapowsky, "Facebook Exposed 87 Million Users to Cambridge Analytica," Wired, [Online]. Available: <https://www.wired.com/story/facebook-exposed-87-million-users-to-cambridge-analytica/>.
- [7] K. Kopp, P. U. Riechert and C. Wolfie, "Corporate surveillance in everyday life," *Cracked Labs*, 2017.
- [8] Agenda digitale EU, "Il datagate dieci anni dopo," [Online]. Available: <https://www.agendadigitale.eu/sicurezza/privacy/il-datagate-dieci-anni-dopo-cosi-ha-cambiato-i-concetti-di-privacy-e-sorveglianza/>.
- [9] G. Vyas, "Data Broker Market Research Report Information," Market Research Future, 2024. [Online]. Available: [https://www.marketresearchfuture.com/reports/data-broker-market-11676#:~:text=The%20Data%20Broker%20Market%20segmentation%2C%20based%20on%20Data%20Type%20includes,market%20revenue%20\(142.5%20Billion\)..](https://www.marketresearchfuture.com/reports/data-broker-market-11676#:~:text=The%20Data%20Broker%20Market%20segmentation%2C%20based%20on%20Data%20Type%20includes,market%20revenue%20(142.5%20Billion)..)
- [10] Experian, "Business model," [Online]. Available: <https://www.experianplc.com/what-we-do/our-business-model>.
- [11] Acxiom, "Services," [Online]. Available: <https://www.acxiom.co.uk/what-we-do/services/>.
- [12] Y. Gu, L. Madio and Reggiani Carlo, "Data Brokers Co-Opetition," 2021.
- [13] Matrice Digitale, "Il governo degli Stati Uniti acquista dati personali dai broker di dati commerciali," 2023. [Online]. Available: <https://www.matricedigitale.it/sicurezza-informatica/il-governo-degli-stati-uniti-acquista-dati-personali-dai-broker-di-dati-commerciali/>.
- [14] Pawtocol, "How much is user data worth?," 2020. [Online]. Available: <https://pawtocol.medium.com/how-much-is-user-data-worth-f2b1b0432136>.
- [15] J. Sadowski, "Companies are making money from our personal data – but at what cost?," The Guardian, 2016. [Online]. Available:



**Politecnico
di Torino**

<https://www.theguardian.com/technology/2016/aug/31/personal-data-corporate-use-google-amazon>.

- [16] TelecomLead, "The data brokering industry is projected to rise to \$365 bn by 2029," 2023. [Online]. Available: <https://telecomlead.com/broadband/the-data-brokering-industry-is-projected-to-rise-to-365-bn-by-2029-109313>.
- [17] "Global Data Broker Market Size, Share, Opportunities, COVID-19 Impact, And Trends By Data Type (Consumer Data, Business Data), By End-User (BFSI, Retail, Automotive, Construction, Others), And By Geography - Forecasts From 2023 To 2028".
- [18] WebFX, "What are data brokers?," [Online]. Available: <https://www.webfx.com/blog/internet/what-are-data-brokers-and-what-is-your-data-worth-infographic/>.
- [19] Transparency market research, "Data brokers market research," 2022. [Online]. Available: <https://www.transparencymarketresearch.com/data-brokers-market.html>.
- [20] R. Reitman, "A Deep Dive into Facebook and Datalogix: What's Actually Getting Shared and How You Can Opt Out," Electronic Frontier Foundation, 2012. [Online]. Available: <https://www.eff.org/deeplinks/2012/09/deep-dive-facebook-and-datalogix-whats-actually-getting-shared-and-how-you-can-opt>.
- [21] A. Rieke, H. Yu, D. Robinson and J. van Hoboken, "Data Brokers in an Open Society," University of Amsterdam, 2016.
- [22] S. Dharmadhikari, "Europe Data Broker Services Market Report 2024," 2024. [Online]. Available: <https://www.cognitivemarketresearch.com/data-broker-services-market-report?srsId=AfmBOoql5hMhZpBfXGEWMXAS7m3RAGymIbZxcFK9BRqxlZfj-98gpjVr>.
- [23] A. Eser, "Global Data Broker Industry Statistics: Revenue Soars, Marketers Embrace Data.," Worldmetrics.org, 2024. [Online]. Available: <https://worldmetrics.org/data-broker-industry-statistics/>.
- [24] "Experian - M&A Summary and Business Overview," Mergr.com, [Online]. Available: <https://mergr.com/company/experian>.
- [25] "Equifax - M&A Summary and Business Overview," Mergr.com, [Online]. Available: <https://mergr.com/company/equifax>.
- [26] A. Kuempel, "The Invisible Middleman: A Critique and Call for Reform of the Data Broker Industry," *Northwestern journal of international law and business*, 2017.
- [27] Consiglio Europeo, "Normativa sui dati: il Consiglio e il Parlamento raggiungono un accordo sull'accesso equo ai dati e sul loro utilizzo," 2023. [Online]. Available: <https://www.consilium.europa.eu/it/press/press-releases/2023/06/27/data-act-council-and-parliament-strike-a-deal-on-fair-access-to-and-use-of-data/>.
- [28] Agenda Digitale, "Data Act: come le aziende dovranno adeguarsi," 2023. [Online]. Available: <https://www.agendadigitale.eu/mercati-digitali/data-act-come-le-aziende-dovranno-adeguarsi-alle-nuove-normative/>.



**Politecnico
di Torino**

- [29] Agenda Digitale, "IA e concorrenza, rebus difficile da risolvere: le sfide per le Autorità," [Online]. Available: <https://www.agendadigitale.eu/mercati-digitali/ia-e-concorrenza-rebus-difficile-da-risolvere-le-sfide-per-le-autorita/>.
- [30] Encyclopedia of Arkansas, "Acxiom," [Online]. Available: <https://encyclopediaofarkansas.net/entries/acxiom-2155/>.
- [31] LiveRamp, "Fiscal results," [Online]. Available: <https://investors.liveramp.com/news-releases/news-release-details/liveramp-announces-fourth-quarter-and-fiscal-year-results-42>.
- [32] Acxiom, "Customer Data," [Online]. Available: <https://www.acxiom.com/customer-data/>.
- [33] redresscompliance, "What is Oracle? The software juggernaut history," 2024. [Online]. Available: <https://redresscompliance.com/what-is-oracle-the-software-juggernaut-history/>.
- [34] M. Hall, "Oracle Corporation," *Encyclopedia Britannica*, 2024.
- [35] Oracle, "Financials," 2023. [Online]. Available: https://s23.q4cdn.com/440135859/files/doc_financials/2023/q4/4q23-pressrelease-June-final.pdf.
- [36] Wikipedia, "Experian," [Online]. Available: <https://en.wikipedia.org/wiki/Experian>.
- [37] Experian, "Financial statement," 2022. [Online]. Available: <https://www.experianplc.com/content/dam/marketing/global/plc/en/assets/documents/reports/2022/annual-report/reports/101-notes-to-the-company-financial-statements-1034-T01.html>.
- [38] Experian, "Financials," [Online]. Available: <https://www.experianplc.com/content/dam/marketing/global/plc/en/assets/docudocum/results-and-presentations/2024/experian-fy24-full-year-results-presentation.pdf>.
- [39] Epsilon, "About us," [Online]. Available: <https://www.epsilon.com/us/about-us>.
- [40] Growjo, "Epsilon revenue and competitors," [Online]. Available: <https://growjo.com/company/Epsilon>.
- [41] Wikipedia, "CoreLogic," [Online]. Available: <https://en.wikipedia.org/wiki/CoreLogic>.
- [42] Corelogic, "Financials," [Online]. Available: https://last10k.com/sec-filings/clgx#link_fullReport.
- [43] Agenda digitale EU, "Data broker, chi sono e come evitano di violare la privacy," [Online]. Available: <https://www.agendadigitale.eu/cittadinanza-digitale/chi-sono-i-data-broker-e-come-evitare-di-violare-la-privacy/#:~:text=Il%20data%20broker%20che%20tratti,proporzionalit%C3%A0%20del%20trattamento%20in%20relazione>.
- [44] Committee on commerce, science, and transportation, "A Review of the Data Broker Industry: Collection, Use, and Sale of Consumer Data for Marketing Purposes," 2013.
- [45] Commissione Europea, "GDPR," 2016.
- [46] Agenda digitale EU, "I data broker e il vero prezzo dei nostri dati: che c'è da sapere," 2018. [Online]. Available: <https://www.agendadigitale.eu/sicurezza/privacy/i-data-broker-e-il-vero-prezzo-dei-nostri-dati-che-ce-da-sapere/>.



**Politecnico
di Torino**

- [47] "Why Perfect Competition Is Desirable," [Online]. Available: https://saylordotorg.github.io/text_principles-of-managerial-economics/s06-07-why-perfect-competition-is-des.html.
- [48] J. Kramer, D. Schnurr and B. Haberer, "Do Consumers Benefit from Selling Their Data? The Economic Impact of Personal Data Brokers on Digital Markets".
- [49] H. Koski, "How Do Competition Policy and Data Brokers Shape Product Market Competition?".
- [50] L. Kaplan, A. Mislove and P. Sapiezynski, "Measuring Biases in a Data Broker's Coverage," 2017.
- [51] H. Ruschemeier, "Data Brokers and European Digital Legislation".
- [52] G. Anthes, "Data brokers are watching you, Communications of the ACM," 2015.
- [53] Malwarebytes, "Broker di dati," Malwarebytes, [Online]. Available: <https://www.malwarebytes.com/it/cybersecurity/basics/data-brokers>.
- [54] M. Favaretto, E. De Clercq and B. S. Elger, "Big Data and discrimination: perils, promises and solutions. A systematic review".
- [55] K. Caitlin, "The Equifax Data Breach and the Resulting Legal Recourse".
- [56] TrendMicro, "When Data Pours: Exactis Database Leaks 340 Million PI; Fastbooking, Ticketmaster UK Breached," [Online]. Available: <https://www.trendmicro.com/vinfo/it/security/news/cybercrime-and-digital-threats/when-data-pours-exactis-database-leaks-340-million-pi-fastbooking-ticketmaster-uk-breached>.
- [57] Twingate, "Verifications.io Data Breach: What & How It Happened?," [Online]. Available: <https://www.twingate.com/blog/tips/verifications-io-data-breach>.
- [58] Theweek.com, "A controversial data firm is selling abortion clinic visitors' location data," [Online]. Available: <https://theweek.com/roe-v-wade/1013221/a-controversial-data-firm-is-selling-abortion-clinic-visitors-location-data>.
- [59] Microsoft, "National Public Data breach: What you need to know," [Online]. Available: <https://support.microsoft.com/en-us/topic/national-public-data-breach-what-you-need-to-know-843686f7-06e2-4e91-8a3f-ae30b7213535>.
- [60] Experian, "Take better control of your identity. Start for free today.," [Online]. Available: <https://www.experian.com/protection/identity-theft-and-credit-protection/>.
- [61] M. Herland, T. M. Khoshgoftaar and R. A. Bauder, "Big Data fraud detection using multiple medicare data sources".
- [62] Agenda digitale EU, "Data Governance Act ora applicativo: così cambia l'economia digitale," 2023. [Online]. Available: <https://www.agendadigitale.eu/sicurezza/privacy/la-data-economy-alla-prova-del-data-governance-act-lo-scenario/>.
- [63] Commissione Europea, "Data Governance Act".
- [64] Agenda Digitale dell'Emilia Romagna, "Digital Services Act e Digital Markets Act," [Online]. Available: <https://digitale.regione.emilia-romagna.it/notizie/archivio/2023/giugno/digital-services-act-e-digital-markets-act-le-normative-dell2019unione-europea-sui-servizi-e-sul-mercato-digitale>.
- [65] Journal of the European Union, "Digital Services Act".



**Politecnico
di Torino**

- [66] Journal of the European Union, "Digital Markets Act".
- [67] USA Government, "Gramm-Leach Bliley Act," 1999.
- [68] Federal Trade Commission, "Incorporating U.S. SAFE WEB Act amendments of 2006," 2006.
- [69] State of California, "California Consumer Privacy Act".
- [70] State of California, "California Privacy Rights Act".
- [71] Smarsh, "The Quest for a Comprehensive Federal Privacy Law: ADPPA vs APRA," [Online]. Available: https://www.smarsh.com/blog/thought-leadership/the-quest-for-a-comprehensive-federal-privacy-law-adppa-vs-apra?utm_source=chatgpt.com.
- [72] USA Congress, [Online]. Available: <https://crsreports.congress.gov/product/pdf/LSB/LSB11161>.
- [73] R. D. Bruin, "A Comparative Analysis of the EU and U.S. Data Privacy Regimes A Comparative Analysis of the EU and U.S. Data Privacy Regimes and the Potential for Convergence".
- [74] K. Weber, "ACCC investigates Australian data broker practices," [Online]. Available: ACCC investigates Australian data broker practices - Strategy - Storage - iTnews.
- [75] CHOICE, "Digital platform services inquiry – march 2024 report on data brokers".
- [76] Deloitte, "The China Personal Information Protection Law (PIPL)".
- [77] People's Republic of China, "Personal Information Protection Law of the People's Republic of China".

Tabella delle figure

Profilazione dei consumatori [7]	17
Scopi dei dati venduti [1]	19
Crescita del mercato dei data broker [14] [15] [16] [17] [18] [19]	24
Suddivisione del mercato in Europa [22]	26



**Politecnico
di Torino**

Grafico di suddivisione surplus all'equilibrio [57]	38
Grafico di suddivisione surplus non all'equilibrio [57]	39
Grafico di accuratezza dei dati in base a discriminante etnica [66]	43