POLITECNICO DI TORINO

Corso di Laurea Magistrale in Ingegneria Gestionale

Tesi di laurea magistrale

# ECONOMIA DEI BIG DATA

Relatore:
Prof. Carlo Cambini

Candidato:
Michele Ingrassia

Anno Accademico 2019-2020

*Alla mia Famiglia…*

# Sommario

**INTRODUZIONE ..... 6**

**1. I Big Data ..... 7**

1.1 Definizione ..... 7

1.2 Qualità dei dati ..... 8

1.3 Diversi approcci di catalogazione dei dati ..... 8

1.4 Le 3V dei Big Data: quali sono? ..... 9

1.5 Perché oggi parliamo di 5V? ..... 10

1.6 La catena del valore ..... 11

1.7 Big data: qual è il loro valore? ..... 13

1.8 L'economia digitale ..... 14

1.9 I Big Data nei vari settori ..... 16

1.10 Professioni nel digitale ..... 18

**2. Piattaforme digitali ..... 20**

2.1 Definizione ..... 20

2.2 Il ruolo delle esternalità ..... 20

2.3 Prezzi nei mercati a due versanti ..... 21

2.3.1 È più efficiente avere due imprese distinte o una piattaforma singola? ..... 22

2.4 I big data e i prezzi personalizzati ..... 26

2.4.1 Gradi di discriminazione ..... 26

2.4.2 Impatto prezzi personalizzati su welfare ..... 28

2.5 I big data e la pubblicità personalizzata ..... 29

2.5.1 Targetability ..... 30

2.5.2 Measurability ..... 31

**3. CONCENTRAZIONE DEI MERCATI ..... 32**

3.1 Switching cost come fonte di vantaggio per l'operatore storico? ..... 32

3.2 I big data e il potere di mercato ..... 33

3.3 Big data come fonte di vantaggio competitivo ..... 34

3.4 Big Data nei mercati a valle ..... 35
3.5 Big data nei mercati a monte ..... 36
3.6 Aspetto comportamentale: effetti di rete ..... 36
3.7 Ignoranza del consumatore e asimmetria informativa ..... 40
3.8 I big data tra monopoli naturali? ..... 41
**4 Il mercato delle APP .....42**
4.1 Il ruolo degli app store..... 42
4.2 Il ruolo delle esternalità di rete e il mercato degli App store ..... 45
4.3 Concentrazione nel mercato delle app ..... 47
4.4 Una soluzione di mercato alle transazioni di dati: i permessi ..... 48
4.4.1 Correlazione tra concentrazione del mercato e permessi..... 50
4.4.2 Correlazione download-permessi ..... 52
4.4.3 Correlazione prezzi-permessi..... 53
**5. Privacy .....56**
5.1 Definizione ..... 56
5.2 Dipendenza dal contesto..... 59
5.3 Malleabilità e influenza ..... 61
5.4 Paradosso della privacy ..... 62
5.5 Monetising privacy ..... 63
5.5.1 Valore del dato come disponibilità ad accettare degli incentivi monetari..... 63
5.5.2 Valore del dato come disponibilità a pagare ..... 64
5.5.3 Valore del dato oggettivo..... 67
5.5.4 Confronto tra i tre valori ..... 68
**6. GDPR .....72**
6.1 Le principali novità introdotte dal GDPR..... 73
6.1.1 Applicazione territoriale..... 74
6.1.2 Consenso ..... 74
6.1.3 Responsabilità congiunta ..... 74
6.2 Efficacia GDPR nel breve e nel lungo periodo ..... 75
6.3 Conseguenze sulla struttura del mercato e sulla concorrenza. ..... 78

**CONCLUSIONI .....81**

**BIBLIOGRAFIA .....84**

**SITOGRAFIA .....87**

**RINGRAZIAMENTI .....88**

# INTRODUZIONE

Il presente lavoro di tesi si pone l'obiettivo di descrivere le principali caratteristiche dei mercati digitali, soffermandosi in particolar modo sul valore dei Big Data.

Nel primo capitolo, viene affrontato il tema generale dei Big Data, evidenziando e descrivendo i principali elementi che li caratterizzano. Si sottolineano, inoltre, le varie fasi della catena del valore dei dati, che consentono di identificare i vari passaggi attraverso cui generare valore, e si descrivono le principali applicazioni dei big data nei vari settori.

Nel secondo capitolo, sono introdotte le piattaforme digitali, che ricoprono un ruolo fondamentale nei mercati digitali. Si osservano i meccanismi di mercato che li caratterizzano, soffermandosi in particolar modo sul concetto di mercato a due versanti e, di conseguenza, sulle possibili strategie di prezzo e sull'utilizzo dei Big data nelle pratiche di discriminazione di prezzo e nella pubblicità personalizzata.

I mercati digitali sono spesso caratterizzati da dinamiche winner take all, dove il vincitore si prende tutto il mercato. Vi è, quindi, un'elevata concentrazione di mercato, una dispersione dei profitti maggiore e forti divari economici tra le imprese.

Nel terzo capitolo, si cerca di scavare più a fondo sulle possibili cause che favoriscono queste dinamiche di mercato, ponendo attenzione maggiore sul binomio Big Data-Potere di mercato**.**

Nel quarto capitolo, si esamina attentamente il mercato delle app. Uno dei principali meccanismi attraverso cui i consumatori cedono i propri dati avviene, infatti, tramite il *download* e il successivo utilizzo delle app. Si analizzano il ruolo degli App store, le dinamiche concorrenziali all'interno del mercato e il tema dei permessi legato alla tutela dei dati più sensibili.

Sulla falsa riga dell'ultimo paragrafo del quarto capitolo, nel quinto capitolo si affronta, più da vicino, uno dei problemi più urgenti e rilevanti legati all'utilizzo dei Big Data dei nostri tempi: la privacy. Si esamina con molta attenzione i temi del paradosso della privacy e della monetizzazione della privacy, esponendo quanto la ricerca è riuscita a far emergere finora.

Nel sesto e ultimo capitolo viene presentato il GDPR, analizzando, in particolar modo, la sua efficacia nel breve e nel lungo periodo e le sue conseguenze sulla struttura e sulla concorrenza nei mercati digitali.

# 1. I Big Data

## 1.1 Definizione

I big data sono una delle evoluzioni più pervasive e profonde del mondo digitale.

Non esiste una definizione univoca di Big Data: questo termine indica un fenomeno astratto, complesso, nato a partire dagli anni Duemila, grazie all'enorme quantità di dati e dataset disponibili e grazie ai progressi tecnologici.

I settori inizialmente interessati dall'incremento di informazioni sono stati la genomica e l'astronomia, ma la diffusione è stata rapida anche in altri settori, dal business alla medicina.

Esistono diverse definizioni che differiscono per alcuni elementi.

"Nel 2011, la McKinsey and Company (multinazionale di consulenza aziendale) ha definito il fenomeno come la nuova frontiera dell'innovazione, della concorrenza e della produttività". Secondo la multinazionale, i big data sono, infatti, quel set di dati che non possono essere acquisiti e gestiti da database classici.

I dati e le tecnologie di analisi, soprattutto nell'ultimo decennio, hanno assunto il carattere di risorsa indispensabile per la gestione aziendale.

I Big Data nascono grazie allo sviluppo tecnologico che ha portato un aumento delle risorse di gestione dei dati: ci sono 4,6 miliardi di cellulari in tutto il mondo e circa 2 miliardi di persone che possono accedere ad internet.

La capacità effettiva mondiale di scambiare informazioni attraverso le reti di telecomunicazione è quasi raddoppiata ogni 40 mesi dal 1989.

Nel 2010 Eric Schmidt sosteneva: "Nel giro di due giorni produciamo la stessa quantità di dati generati dagli albori della civiltà sino al 2003".

Secondo IBM: "Entro il 2025 il mondo avrà generato una massa di dati dell'ordine dei 160 zettabyte". Nel mondo, in ogni minuto, sono scambiate milioni di informazioni. Ogni 60 secondi, su Facebook, sono creati 3,3 milioni di post, pubblicati 510.000 commenti e aggiornati 293.000 stati, su WhatsApp sono scambiati 29 milioni di messaggi e su Google sono effettuate 3,8 milioni di ricerche.

È evidente, quindi, come gli utenti si rapportano con un flusso continuo e ininterrotto di informazioni e come, quindi, il mondo odierno sia sommerso da un enorme quantità di dati.

## 1.2 Qualità dei dati

La qualità dei dati è determinata da un insieme di caratteristiche:

- **Completezza:** presenza di tutti i dati necessari per descrivere un'entità, un evento o una transazione.

- **Consistenza:** assenza di contraddizione nei dati. La consistenza è legata al concetto di equivalenza delle informazioni. Per esempio, in una banca, la somma del saldo di fine mese precedente di un conto corrente con i movimenti attivi e passivi, deve essere uguale al saldo di fine mese corrente.

- **Accuratezza**: riguarda la conformità ai valori reali, cioè la correttezza dei valori stessi. Il punto di partenza per l'accuratezza dei dati risiede nella loro integrità. Oltre alla necessità di una doppia manutenzione, la duplicazione dei dati incide in maniera negativa sulla qualità, perché è possibile una mancata sincronizzazione tra le copie dei dati.

## 1.3 Diversi approcci di catalogazione dei dati

È possibile, inoltre, distinguere diversi approcci di catalogazione dei dati, con riferimento alla:

- **Tipologia di informazione:** i dati possono essere classificati in base alle diverse tipologie di informazioni che sono in grado di fornire. I dati possono, ad esempio, fornire informazioni su individui, sui loro comportamenti, le loro preferenze e la loro posizione geografica.

- **Struttura del dato:** i dati possono essere differenziati a seconda se siano strutturati o meno. Da questa particolarità, dipende la possibilità di estrarne valore economico. Per dati strutturati, cioè dati numerici o di natura finanziaria (tabelle, documentazioni di *office automation*, dati M2M), è necessario un modello di estrazione che definisca una serie di parametri come il settore di riferimento, la tipologia di dati che interessano quel dato settore e infine la relazione esistente tra i dati stessi.

Un esempio di dati strutturati potrebbe essere rappresentato da una base di dati di un consumatore che contiene informazioni relative al nome, cognome, indirizzo, età, recapito telefonico, ecc.

I dati strutturati possono essere più facilmente elaborati e utilizzati per scopi commerciali rispetto a dati destrutturati.

Ai dati destrutturati, ossia tutti quei dati che non rientrano in categorie definite (testi di e-mail, dati video e audio, dati social, immagini), viceversa, non corrisponde un modello specifico di elaborazione e generalmente necessitano di essere trattati da diversi e più innovativi algoritmi affinché acquisiscano un determinato valore commerciale.

- **Modalità di raccolta:** i dati possono essere raccolti attraverso diverse modalità e a seconda della modalità adottata si rileva un diverso impatto sul loro valore economico.

  I dati sono, di solito, forniti volontariamente ed attivamente dagli utenti. Ad esempio, per effettuare acquisti online, l'utente deve rilasciare tutta una serie di informazioni personali, tra cui indirizzo, dati di pagamento ed e-mail o recapiti telefonici per perfezionare l'acquisto. I social network e le applicazioni di messaggistica istantanea come WhatsApp e Telegram si basano essenzialmente sui dati personali forniti dai loro utenti come nome, indirizzo, sfondo educativo e messaggi personali, foto, video, commenti sulle recenti notizie, le preferenze commerciali ecc.

## 1.4 Le 3V dei Big Data: quali sono?

Nel 2001, Doug Laney, allora vicepresidente e Service Director dell'azienda Meta Group, descrisse in un report il *Modello delle 3V* relativo alle 3V dei Big Data:

- **Volume**: è certamente la proprietà che più facilmente si può accostare ai *big data*. Numerosi studi e statistiche hanno cercato di misurare tale proprietà e l'unità di misura più idonea è quella dei *zettabyte*: 1 *zettabyte* corrisponde a una capacità di archiviazione pari a oltre 36.000 anni (in termini di durata) di video in HD.

- **Velocità:** i dati si originano e si raccolgono sempre più rapidamente, grazie anche all'utilizzo di dispositivi dotati di sensoristica.
  Tuttavia, le aziende non devono solo raccogliere le informazioni velocemente, ma devono anche analizzarle rapidamente, in tempo reale, per poter prendere decisioni di business più tempestivamente possibile.

- **Varietà**: "*More isn't just more. More is different.*" – così scriveva, nel 2008, Chris Anderson sul magazine Wired.
  Con varietà si fa riferimento, proprio, alle diverse tipologie di dati oggi utilizzabili, che provengono da tantissime fonti anch'esse differenti.

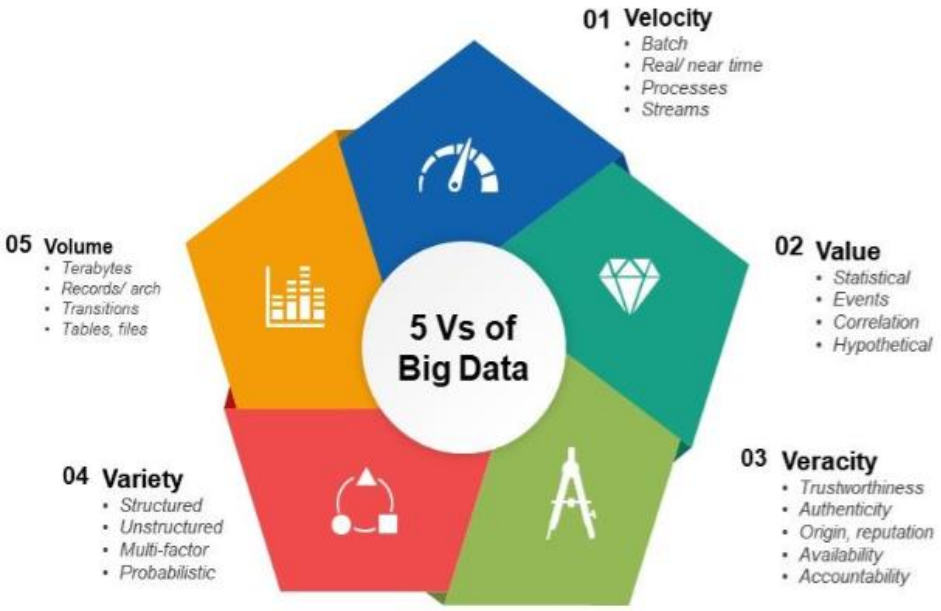
## 1.5 Perché oggi parliamo di 5V?

Nei primi anni 2000 si definivano, quindi, i Big Data con tre parole: volume, velocità e varietà. Con il passare degli anni, mentre il termine diveniva sempre più concreto, ci si è chiesti se non vi fossero altre caratteristiche da mettere in risalto. Oggi il modello di Laney è stato impreziosito da altre due variabili, che definiscono come questi nuovi dati debbano essere utilizzati:

- **Veridicità:** alcuni usano dire *"Bad data is worse than no data"*. I dati devono, infatti, essere affidabili e raccontare la verità.
  Con i Big Data, questa sfida è ancora più complicata da sostenere: le tecnologie di gestione dei dati cambiano come cambia anche la rapidità con la quale i dati sono raccolti e aumentano le fonti. Non è facile, quindi, mantenere l'integrità e la qualità delle informazioni che rimangono un pilastro imprescindibile per dar vita ad analisi che siano affidabili e utili.

- **Valore**: I Big Data sono stati definiti, negli ultimi anni, come il nuovo petrolio o il nuovo oro, cioè una sorgente piena di valore. Attraverso il loro uso, un'impresa può estrarre valore, prendere decisioni più informate, appropriate e veloci.
  Tuttavia, limitarsi a raccogliere i dati, pur sfruttando le migliori tecnologie disponibili sul mercato, non garantisce di avere informazioni e, soprattutto, di estrarre conoscenza. Per attuare questo processo, e far sì che i Big Data possano essere trasformati in informazioni da utilizzare nei processi aziendali costruendo

conoscenza per migliorare le performance e per identificare nuove opportunità commerciali, sono necessari strumenti di Big Data Analytics.

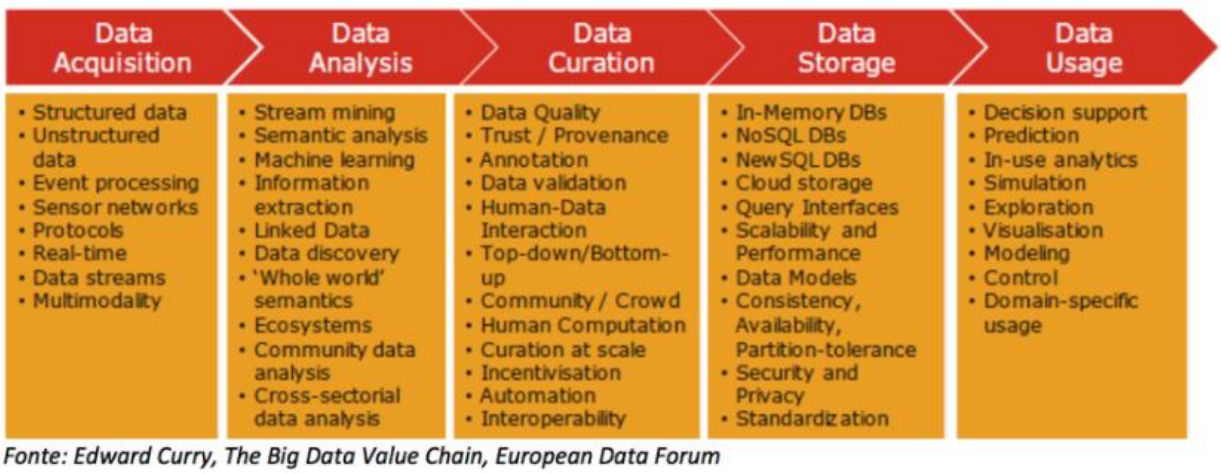



*"Figura 1: le 5 V dei Big Data"*

## 1.6 La catena del valore

La catena del valore descrive il flusso di informazioni attraverso una serie di fasi necessarie, affinché possa essere generato valore dall'insieme di dati considerati.

La Commissione Europea, già nel 2013, ha ritenuto che la catena del valore dei dati **(Figura 2)** fosse il fulcro dell'economia del futuro, con lo scopo di determinare opportunità di sviluppo digitale anche nei settori più tradizionali, quali ad esempio trasporti, servizi finanziari, sanità, manifatturiero, vendita al dettaglio.




*"Figura 2: la catena del valore dei dati"*

In particolare:

- **Acquisizione**: è il processo di raccolta, filtraggio, e selezione dei dati prima del loro inserimento in un archivio su cui l'analisi dei dati può essere eseguita. L'acquisizione dei dati è una delle principali sfide, se consideriamo i requisiti infrastrutturali richiesti. L'infrastruttura, necessaria a supportare la raccolta di dati di grandi dimensioni, deve diminuire la possibile latenza nell'acquisizione dei dati, essere in grado di gestire volumi di transazioni molto elevati e supportare strutture di dati dinamiche e flessibili.

- **Analisi**: comporta l'esplorazione, la trasformazione, e la configurazione dei dati con l'obiettivo di evidenziare quelli più rilevanti e di sintetizzare e ricavare informazioni nascoste ad elevato potenziale dal punto di vista del business.

- **Cura**: è la gestione attiva dei dati, durante il loro ciclo di vita, per garantire il rispetto dei requisiti di qualità necessari per un uso più efficace. Per eludere il rischio che un sistema così diffuso di raccolta dati possa creare un fenomeno di over-collection, cioè di ridondanza del dato, si effettua un secondo step di lavorazione del dato che riguarda la sua preparazione e conservazione per gli utilizzi successivi. È questa la fase in cui il dato inizia a trasformarsi in informazione attraverso capacità elaborative e software che consentono sia una facile scalabilità dei dati, sia una loro corretta conservazione. La cura dei dati è eseguita da esperti professionisti che sono responsabili del miglioramento dell'accesso al dato e della qualità. I curatori devono inoltre garantire che i dati siano affidabili, riutilizzabili, e adattabili agli obiettivi prefissati.

- **Archiviazione**: rappresenta la modalità di gestione dei dati in modo modulare, che soddisfa le esigenze delle applicazioni che richiedono un accesso veloce ai dati. Relational Database Management Systems (RDBMS o modelli relazionali) sono stati la principale e quasi unica soluzione di archiviazione per quasi 40 anni.
  Negli ultimi anni, inoltre, si sta espandendo la tecnologia cloud. Sempre molte più aziende adottano il modello del cloud computing poiché permette di gestire e archiviare l'enorme quantitativo di big data che si produce giorno dopo giorno, migliorando le proprie economie di scala.

Per cloud computing si intende quel modello che consente un comodo accesso ai dati memorizzati. I dati sono condivisi con altri utenti e collocati in modo dinamico.

- **Uso**: si riferisce a quell'insieme di attività di business basate sui dati in cui l'accesso, l'analisi e la predisposizione di strumenti che integrano l'analisi dei dati all'interno dell'attività d'impresa, sono fondamentali.
  Lo stadio finale del dato è quello in cui da semplice informazione esso si trasforma in conoscenza e visione.
  L'utilizzo dei dati nell'ambito dei processi decisionali aziendali può rafforzare la competitività dell'impresa attraverso una diminuzione dei costi e un incremento del valore aggiunto.

Le varie fasi della catena del valore del dato consentono così di modellare il sistema dei big data e, di conseguenza, di identificare i vari passaggi attraverso cui generare valore, e, più in generale, conoscenza dai dati. Sono le competenze e le tecnologie di Analytics a trasformare i dati grezzi in informazioni di valore. È oggi possibile ottenere vantaggio competitivo grazie a decisioni più rapide e più informate, non solo per le organizzazioni più grandi ma anche per le piccole e medie imprese**.**

## 1.7 Big data: qual è il loro valore?

Il valore dei dati nasce, secondo Cukier e Mayer-Schönberger, oltre che da un utilizzo primario per estrarre informazioni e fare delle previsioni, anche dai possibili usi secondari ricavabili se si ricombinano i dati. Il valore secondario deriva dal possibile riutilizzo futuro, cioè, dall'unione di più dataset, che permettono di raccogliere informazioni combinate e arricchite.

Prevedendo trend e tendenze, i Big Data permettono, infatti, grazie a adeguate analisi, di sviluppare nuovi prodotti e servizi e di migliorare quelli già in circolazione, attraverso, ad esempio, servizi post-vendita più completi.

I Big Data costituiscono, inoltre, un ottimo strumento per la segmentazione della popolazione, rendendo ogni azione rivolta ai singoli clienti sempre più specifica.

Tuttavia, il valore dei dati è soggetto a una rapida obsolescenza: essi perdono affidabilità con il passare del tempo, a causa della rapidità con cui mutano i trend. La presenza nei

dataset di dati obsoleti fa perdere valore all'intero set di dati ed è, quindi, necessario utilizzare dei modelli sofisticati che permettano di ripulirli dai dati vecchi.

Non tutti i dati hanno lo stesso valore e non tutti i dati sono facilmente recuperabili, perciò possiamo classificare i dati in quattro categorie in base al loro *potenziale* e alla loro *disponibilità*:

*"Figura 3: customer data strategies"*

- **Losing causes:** i dati non sono disponibili e anche se lo fossero non porterebbero valore, perciò sarebbe inutile investire nella loro collezione (es: Ristoranti).

- **Quick wins:** i dati sono disponibili ma il loro utilizzo non porta grande valore, magari migliora i ricavi ma sicuramente non li raddoppia (es: Supermercati).

- **Trade-offs:** i dati che potenzialmente potrei raccogliere hanno un valore elevato ma non sono facilmente reperibili, non abbiamo disponibilità per quei dati.

- **Imperatives:** i dati sono disponibili e hanno un grande valore, è fondamentale acquisirli e poterli utilizzare. Occorre investire in sistemi che permettono di estrarre informazioni utili. Google si trova sicuramente in questa posizione.

## 1.8 L'economia digitale

L'economia digitale è rappresentata da un ecosistema di diverse tipologie di operatori del mercato, come produttori, ricercatori e fornitori di infrastrutture, che collaborano fra loro per rendere i dati accessibili e utilizzabili.

Lo sviluppo di internet ha caratterizzato, nel corso degli ultimi due decenni, un inarrestabile processo di innovazione con una diffusione netta e rapida della digital economy in tutti i settori.

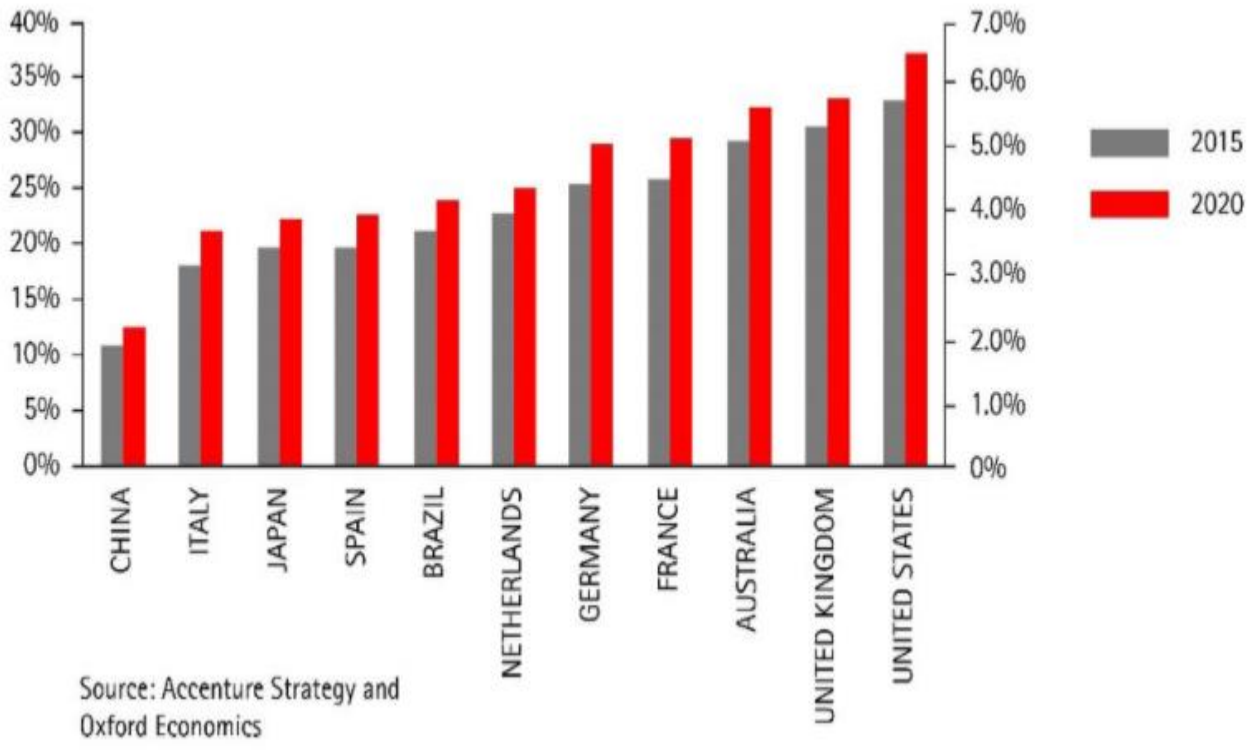
Si stima che i principali benefici per le imprese si avranno in termini di:

- aumento annuale dei ricavi digitali, in media fino al 2,9%;
- riduzione dei costi, in media del 3,6% l'anno;
- forte aumento degli investimenti

La combinazione "intelligente" di competenze digitali, tecnologie come cloud computing, IOT, big data analytics, artificial intelligence e 5G, prendendo come riferimento l'esempio americano, potrebbe far registrare un aumento del PIL del 2,1%.

La **Figura 4** mostra la quota "digitale" del PIL rispetto a 11 paesi considerati.

Per la maggior parte delle economie di riferimento, la quota digitale del PIL cresce di circa il 3% tra il 2015 e il 2020, l'equivalente di un 12,5% di aumento a livello globale. Attualmente, gli Stati Uniti guidano la classifica, con un livello di economia digitale del valore di circa 5.9 miliardi di dollari, pari al 33% del PIL prodotto.




*"Figura 4: quota digitale del PIL paese per paese (2015 e 2020)"*

Il valore complessivo dell'economia dei dati in Europa è passato dai 257 miliardi di euro nel 2014, pari all'1,85% del pil europeo, a 272 miliardi di euro nel 2015, pari all'1,87% del pil dell'UE, a quasi 300 miliardi di euro nel 2016, quasi 2% del PIL europeo.
Nel 2020 la data economy europea è aumentata a 739 miliardi di euro con un impatto complessivo del 4% sul PIL, in uno scenario di crescita elevata, caratterizzato da un ruolo più forte dell'innovazione digitale e con maggiori investimenti ICT complessivi.
Il progetto delineato dalla presidente della Commissione europea, Ursula von der Leyen intende rafforzare sempre più la competitività del vecchio Continente, facendo recuperare il terreno perso in ambito tecnologico rispetto a paesi più "digitali" come Stati Uniti o Cina. Entro il 2025, il valore dell'economia dei dati arriverà a 829 miliardi di euro, pari a circa il 5,8% del pil comunitario. Bruxelles intende, quindi, mettere a disposizione tra i 4 e i 6 miliardi di euro da investire in spazi comuni europei e in una federazione di infrastrutture e servizi cloud.

## 1.9 I Big Data nei vari settori

I big data sono ormai analizzati e usati in moltissimi campi di applicazione: dalla sanità ospedaliera alla cura a distanza, dall'istruzione ai trasporti, dall'informazione giornalistica alla fruizione di contenuti multimediali, dai servizi finanziari e assicurativi alla ricerca del lavoro, dal controllo del traffico stradale alle previsioni del tempo, dall'agricoltura alla domotica e all'IOT e così via.
Per le aziende, le analisi di dati strutturati e non strutturati consentono il generarsi di modelli in grado di analizzare comportamenti d'acquisto, opinioni nei confronti dei prodotti, dell'azienda o dei suoi competitor.

Nel settore bancario e finanziario, l'estrazione ed elaborazione di big data aiuta a individuare transazioni fraudolente, a sviluppare analisi delle tendenze commerciali delle imprese, ma anche ad offrire incentivi personalizzati all'uso di carte di credito in funzione dei profili di reddito e di spesa dei clienti.

In molte applicazioni che riguardano il mercato assicurativo, i dati sono già utilizzati per aiutare gli assicuratori a fissare premi assicurativi in maniera più precisa, identificare richieste fraudolente e migliorare le strategie di marketing. Diverse aziende assicuratrici raccolgono dati relativi alla sicurezza della guida, offrendo sconti agli assicurati che si

prestano ad un monitoraggio della guida, attraverso app scaricate volontariamente su smartphone o altri meccanismi volontariamente installati nell'autoveicolo. Questo sistema di premi per i comportamenti virtuosi comporta la collezione di tantissimi dati.

Anche la grande distribuzione con l'avvento dei Big Data è destinata a un cambiamento. Le telecamere digitali sono usate per monitorare i livelli delle scorte e individuare i consumatori con meccanismi di riconoscimento facciale. Possono essere, inoltre, collegate con i cellulari dei clienti per raccogliere dati sui tempi di permanenza, sui prodotti visionati e acquistati.
A Seattle e a Chicago, ad esempio, sono stati aperti i supermercati "Amazon Go", i primi senza casse all'uscita. Una volta scaricata l'applicazione di Amazon Go, basta loggarsi e accedere al supermercato grazie al codice QR generato automaticamente. Un sistema di fotocamere, sensori e algoritmi di machine learning rileva gli articoli rimossi dagli scaffali e ci addebita automaticamente sul conto, una volta usciti dal supermercato. Questo sistema non solo rende più veloce e agevole la spesa per il cliente, ma ne acquisisce dati di consumo e può incrociarli con le spese medie dell'utente, per prezzi e tipologia, permettendo al supermercato di agevolare il tempo di spesa del consumatore, di gestire al meglio le scorte, ma anche alla piattaforma di e-commerce di Amazon di identificare e profilare meglio quel cliente per fornirgli ulteriori offerte personalizzate.

Anche nel mondo del trasporto, della logistica e della consegna postale di pacchi e merci, le imprese stanno raccogliendo e analizzando i dati dei loro veicoli e usando questi dati per ottimizzare i comportamenti dei conducenti e i percorsi. Sul piano dei trasporti l'utilità della raccolta e dell'elaborazione di dati dispersi è fondamentale per migliorare i servizi e ridurre le varianze relativi ai picchi. Diventa possibile, per ciascuno di noi, organizzare un viaggio articolato su più mezzi di trasporto, ricevendone proposte personalizzate, sia in termini di prezzo che di qualità.

La rivoluzione digitale dei Big Data entra anche nelle nostre case con la domotica e l'Internet delle cose (IoT) e permette di raccogliere dati sui nostri consumi energetici domestici. I big data costruiscono, inoltre, un profilo adeguato della nostra permanenza e dei nostri bisogni e permettono agli elettrodomestici di parlare tra di essi, per ottimizzare decisioni ed evitare sprechi. Le "case intelligenti", con assistenti vocali come quelli offerti da Google e Amazon, raccolgono tantissime informazioni su di noi e per noi.

I benefici non sono però circoscritti unicamente alle imprese: anche governi, agenzie governative e servizi di pubblica sicurezza possono trarre informazioni interessanti dall'analisi dei Big Data.
Un utilizzo efficace del Big Data Analytics, da parte dei governi e delle amministrazioni pubbliche a livello europeo, potrebbe ridurre del 15%-20% i costi amministrativi, grazie a una maggiore efficienza che deriverebbe dalla maggiore trasparenza.
L'impiego dei Big Data da parte dei servizi di intelligence, inoltre, riguarda il monitoraggio automatico di contenuti, in qualsiasi lingua o formato, su siti web e social network, utili per ricercare possibili minacce alla sicurezza del paese.

## 1.10 Professioni nel digitale

Uno dei più grandi ostacoli alla diffusione dei big data è rappresentato dalla scarsità di figure professionali adeguate. Una delle principali sfide dei big data è, quindi, di natura umana e consiste nel diffondere la "cultura del dato" all'interno di tutta l'organizzazione.
Secondo il McKinsey Global Institute, è necessario investire fortemente sulla formazione di figure professionali adatte alla digital economy, e più in generale far aumentare la consapevolezza sul valore dei dati nelle organizzazioni.

Come dimostrano molti studi dell'OCSE sull'offerta e la domanda di nuove professionalità nell'economia digitale, gli impatti organizzativi di big data e algoritmi trasformano i rapporti aziendali, le relazioni all'interno delle aziende e il concetto stesso di lavoro, per come lo abbiamo sempre conosciuto.
Da diversi anni si è inaugurato un lungo periodo di transizione nel quale vecchi mestieri non sono più richiesti dal mercato, mentre nuove professionalità non si sono ancora sviluppate.
L'Agenzia per l'Italia Digitale ha recentemente ribadito:
"In futuro la digitalizzazione renderà necessario un mix più articolato di competenze. Un mix in cui skills di natura tecnologica si uniscono a soft skills, quali pensiero critico, creatività e intelligenza emotiva, capacità di leadership e di gestione del cambiamento.
Le nuove figure come il Digital Strategic Planner, il Digital Manager, il Chief Digital Officer, l'Innovation Manager saranno in grado di contestualizzare e allineare strategie di business ed evoluzioni tecnologiche. Le professioni dovranno combinare competenze tecnologiche

su più fronti (protocolli industriali, IoT, Cloud Computing, Big Data, nuove app, realtà aumentata, robotica e security) e competenze più strategiche".

Un recentissimo studio della Commissione Europea, compiuto sui 28 paesi dell'Unione, ha rivelato che, nel 2015, sei milioni di lavoratori in Europa risultavano occupati all'interno della cosiddetta data industry, raggiungendo 6,16 milioni nel 2016.

Si stima che in uno scenario di crescita elevata, il numero dei data workers in Europa crescerà fino a 12 milioni, ad un tasso di crescita medio complessivo del 15,1% entro il 2022.

Tuttavia, secondo l'analisi, nel 2016 si è verificato un gap importante tra la domanda e l'offerta di lavoratori nella digital economy di ben 420.000 posti di lavoro vacanti, pari al 6,2% della domanda totale dei professionisti di dati specializzati.

# 2. Piattaforme digitali

## 2.1 Definizione

Un ruolo fondamentale nell'economia dei dati, lo ricoprono le piattaforme digitali.

Una piattaforma è un soggetto intermediario che permette la comunicazione tra due gruppi diversi di utenti che, interagendo, ottengono entrambi benefici di rete. L'organizzazione che crea valore, permettendo l'interazione tra i gruppi di utenti, si chiama *piattaforma a più parti* (multi-sided platform).

Le piattaforme emergono in presenza di *frizioni di mercato*, quei fenomeni che rallentano o impediscono uno scambio o una transazione economica tra due agenti, quali la mancanza di informazioni, l'incertezza legata ed esse o dei searching costs elevati.

In un mercato ideale, in assenza di frizioni, le transazioni avverrebbero senza problemi.

## 2.2 Il ruolo delle esternalità

Gli utenti decidono di appoggiarsi ad una piattaforma quando essa permette un'interazione più efficace ed efficiente rispetto all'interazione non mediata. Le piattaforme sono caratterizzate da *esternalità di rete*: il valore che l'utente attribuisce alla piattaforma dipende dal numero complessivo di utenti che ne fanno parte, ovvero dalla dimensione della rete. Un incremento di utenti comporta un incremento di utilità nell'utilizzo della piattaforma.

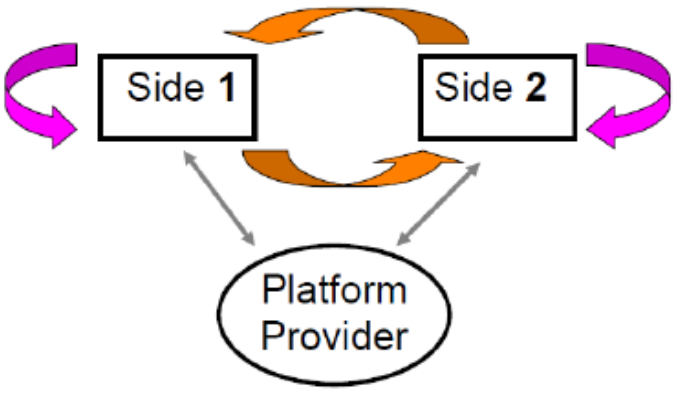
Un mercato a due parti **(Figura 5)** ha generalmente al suo interno quattro effetti di rete che possiamo dividere in due categorie:

- effetti ***same-side*** = l'effetto che il numero di utenti di un gruppo/lato ha sugli utenti dello stesso gruppo/lato.

- **effetti *cross-side*** = l'effetto che il numero di utenti dell'altro gruppo/lato genera sul mio gruppo/lato.

Tali effetti possono essere sia positivi che negativi. Google ad esempio, guadagna molto dall'advertising: crea contenuti per attrarre utenti (side 1) che a loro volta attraggono advertiser (side 2). Il gruppo degli utenti ha un effetto same-side positivo ed un effetto cross-side negativo: maggiori utenti portano a maggiori utenti, maggiori advertiser

possono portare ad una riduzione di utenti. Per quanto, invece, riguarda il gruppo degli advertiser, sono caratterizzati da un effetto same-side negativo e un effetto cross-side positivo: maggiori advertiser generano più competizione all'interno del gruppo, maggiori utenti invece attraggono più advertiser.




*"Figura 5: multi-side Platform"*

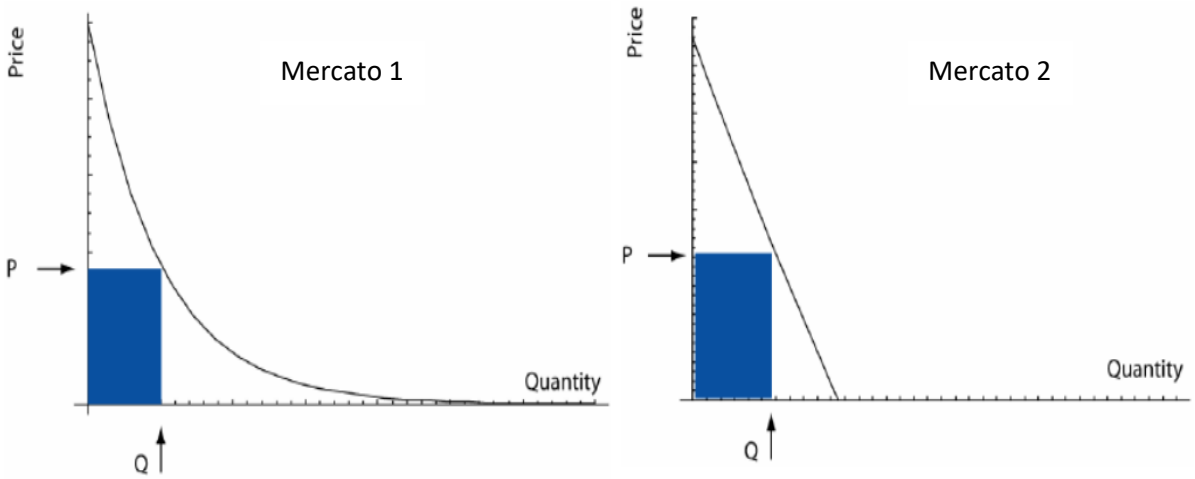
## 2.3 Prezzi nei mercati a due versanti

I recenti studi effettuati da Mark Rysman, G. Parker e M. Van Alstyne hanno affermato che il successo di una piattaforma dipende dalla sua capacità di sfruttare le esternalità.

Il problema nel fissare i prezzi nei due mercati della piattaforma è simile al problema che si ha in presenza di prodotti complementari (es: rasoi/lamette ecc.).

In generale ci sono due possibili soluzioni al problema.

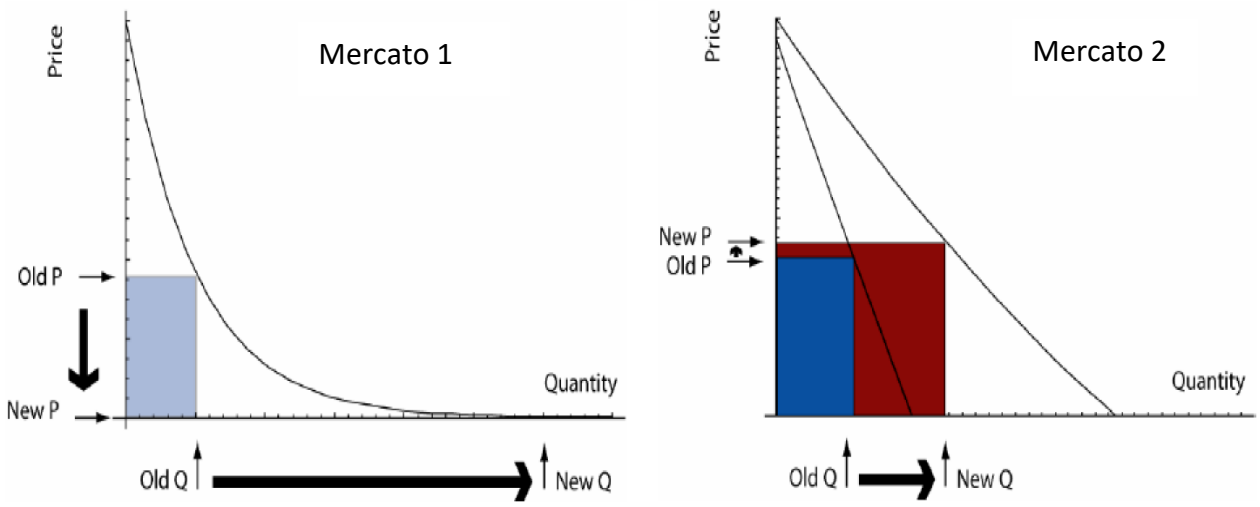
La piattaforma potrebbe fissare due prezzi separati per i due mercati massimizzando le rispettive funzioni di profitto (ricavo marginale = costo marginale). In questo modo la piattaforma avrebbe profitti in entrambi i mercati, come mostrato nella **Figura 6**.




*"Figura 6: prezzi separati in due mercati"*

Una seconda soluzione è data dalla piattaforma che potrebbe ridurre il prezzo o addirittura rendere l'accesso gratis nel primo mercato aumentando così la domanda per il secondo mercato. A causa delle esternalità di rete, una riduzione del prezzo nel primo mercato provoca una crescita della domanda che si riflette anche sul secondo mercato, come mostrato nella **Figura 7**.

Le piattaforme solitamente utilizzano questa seconda metodologia. Facebook, ad esempio, è gratis per gli utenti (Mercato 1), mentre monetizza molto con l'advertising (Mercato 2).




*"Figura 7: prezzi con metodo razor and blade"*

### 2.3.1 È più efficiente avere due imprese distinte o una piattaforma singola?

***Caso 1: due piattaforme distinte***

In questo caso, i due versanti sono serviti da due imprese indipendenti. Si suppone che sul lato 1 agisca l'impresa 1 e che sul lato 2 agisca l'impresa 2.

Si normalizzano a zero i costi di produzione.

Date le funzioni di domanda:

$$D(p_1) = 1 - p_1 \qquad D(p_2) = 1 - p_2$$

$$q_1 = 1 + e_{21}D(p_2) - p_1 \qquad q_2 = 1 + e_{12}D(p_1) - p_2$$

- $e_{12}$= esternalità della domanda 1 sulla domanda 2
- $e_{21}$= esternalità della domanda 2 sulla domanda 1

è possibile massimizzare i profitti delle due imprese sapendo che i profitti sono dati da prezzo per quantità:

$$Max\ p_1 \quad \pi_1 = p_1(1 + e_{21}D(p_2) - p_1)$$

$$Max\ p_2 \quad \pi_2 = p_2(1 + e_{12}D(p_1) - p_2)$$

otteniamo:

$$p_1 = \frac{2 + e_{21}(1 - e_{12})}{4 - e_{12}e_{21}} \qquad \pi_1 = \frac{\left(2 + e_{21}(1 - e_{12})\right)^2}{(4 - e_{12}e_{21})^2}$$

$$p_2 = \frac{2 + e_{12}(1 - e_{21})}{4 - e_{12}e_{21}} \qquad \pi_2 = \frac{\left(2 + e_{12}(1 - e_{21})\right)^2}{(4 - e_{12}e_{21})^2}$$

***Caso 2: unica piattaforma in entrambi i mercati***

L'impresa sceglie i prezzi in modo da massimizzare i propri profitti, dati ora dalla somma dei profitti ottenuti su ciascun versante (profitti congiunti):

$$Max\ p_1p_2 \quad \pi = \pi_1 + \pi_2 = p_1q_1 + p_2q_2$$

Dalle condizioni del primo ordine si ottiene:

$$p_1(p_2) = \frac{1 + e_{21} - p_2(e_{21} + e_{12})}{2}$$

$$p_2(p_1) = \frac{1 + e_{12} - p_1(e_{21} + e_{12})}{2}$$

Mettendo a sistema queste due condizioni, si ricavano i prezzi di equilibrio praticati dall'impresa in ciascun lato del mercato e il profitto di equilibrio:

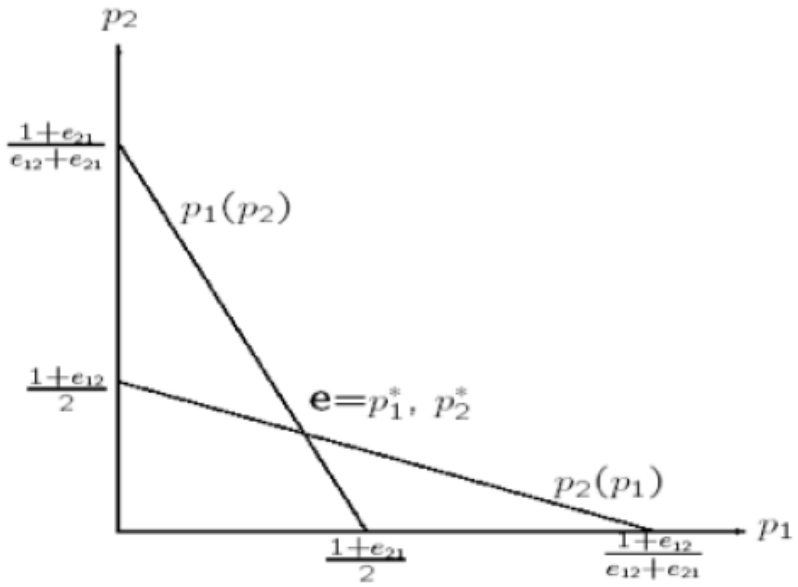
$$p_1' = \frac{1 - e_{12}}{2 - (e_{21} + e_{12})} \qquad p_2{}' = \frac{1 - e_{21}}{2 - (e_{21} + e_{12})}$$

$$\pi' = \frac{1 - e_{12}e_{21}}{2 - (e_{21} + e_{12})}$$

I prezzi di equilibrio dipendono in modo cruciale dall'intensità delle esternalità incrociate. In particolare, si possono verificare tre casi:

- $\boldsymbol{e_{21} = e_{12}}$

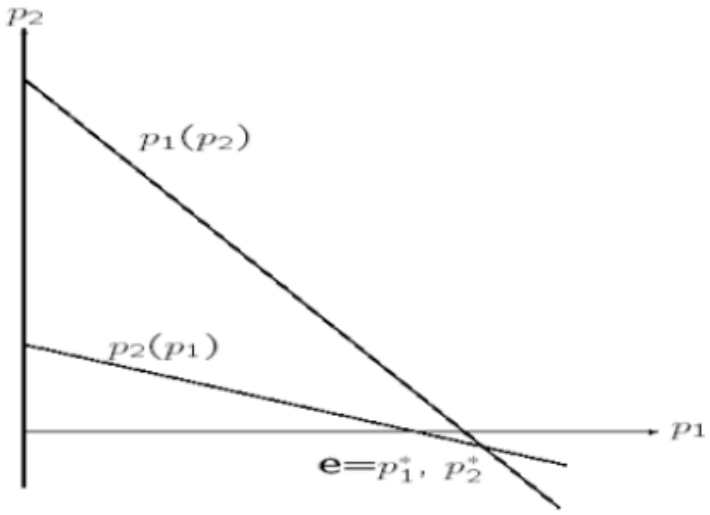
  Non vi è sbilanciamento di esternalità ed entrambi i prezzi sono positivi




*"Figura 8: prezzi di equilibrio con $e_{21} = e_{12}$"*

- $\boldsymbol{e_{21} > e_{12}}$

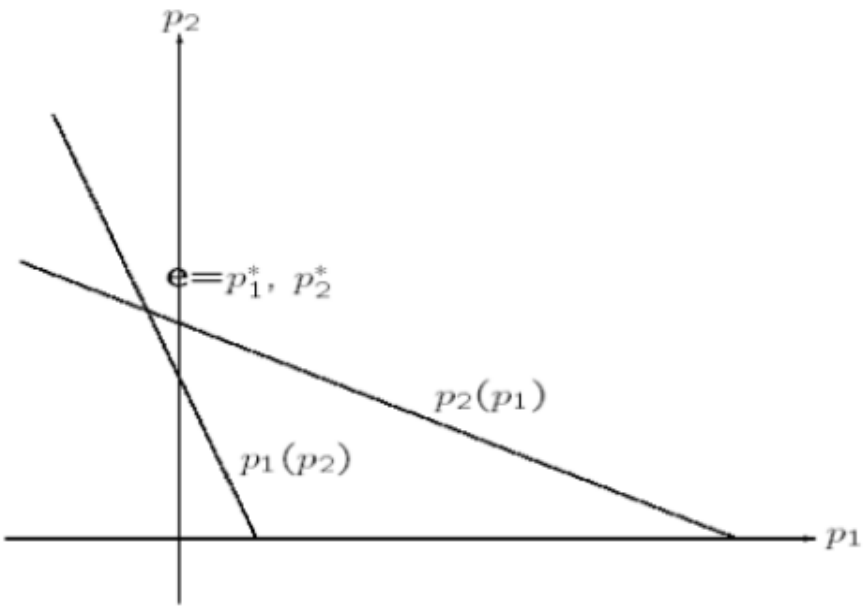
  L'intensità dell'esternalità di rete che agisce dal lato 2 al lato 1 è considerevolmente maggiore. Aumenta, infatti, molto di più la domanda del bene/servizio 1 rispetto a quella del bene/servizio 2. In questo caso il prezzo di equilibrio è "negativo", ossia inferiore al costo di produzione sul lato 2 e positivo sul lato 1. Per tale ragione, è ottimale "sussidiare" il lato 2.




*"Figura 9: prezzi di equilibrio con $e_{21} > e_{12}$"*

- $e_{21} < e_{12}$

  L'Intensità dell'esternalità di rete che agisce dal lato 1 al lato 2 è considerevolmente maggiore. In questo caso il prezzo di equilibrio è "negativo" (ossia inferiore al costo di produzione) sul lato 1 e positivo sul lato 2. È pertanto ottimale "sussidiare" il lato 1.




*"Figura 10: prezzi di equilibrio con $e_{21} < e_{12}$"*

Confrontando i casi 1 e 2, è possibile mostrare che, indipendentemente dal grado di esternalità, è sempre conveniente comportarsi da piattaforma massimizzando i profitti congiunti.

Se le due esternalità sono uguali, i prezzi della piattaforma sono minori rispetto al caso di imprese separate e quindi aumentano il surplus del consumatore, i profitti delle imprese e di conseguenza il benessere collettivo, formando così un equilibrio pareto efficiente.

Nel caso in cui le due esternalità siano diverse, la somma dei prezzi della piattaforma è minore rispetto alla somma dei prezzi di due aziende separate e quindi avere una piattaforma anche in questo caso è più efficiente rispetto ad avere due aziende separate ma non è un equilibrio pareto efficiente.

## 2.4 I big data e i prezzi personalizzati

L'incredibile aumento delle capacità computazionali degli attuali sistemi informatici ha reso le imprese contemporanee capaci di individuare o di approssimare molto bene la disponibilità a pagare dei loro clienti senza dover ricorrere a negoziazioni individuali e senza che tali clienti e consumatori siano informati o siano pienamente consapevoli di questa nuova capacità delle imprese.

Le piattaforme di e-commerce, già da diversi anni, utilizzano algoritmi per la determinazione automatica dei prezzi di vendita. Tali modalità di tariffazione sono ampiamente utilizzate nel settore dei viaggi, prenotazione alberghiere, vendite al dettaglio e nel settore assicurativo.

In economia, l'espressione discriminazione di prezzo viene utilizzata per indicare tutti quei casi di concorrenza imperfetta nei quali un'azienda offre lo stesso bene a prezzi differenti e maggiori del costo marginale di quel prodotto o servizio. Infatti, lo scopo delle aziende che optano per questa strategia è quello di calcolare dei prezzi che, anziché dipendere dai costi di produzione, derivano dalla domanda e, in particolare, dalla disponibilità a pagare (anche detta, prezzo di riserva) dei loro differenti clienti.

Tuttavia, non è semplice individuare il prezzo di riserva dei consumatori o, meglio, non è così comune per un'impresa disporre delle informazioni necessarie per conoscere le preferenze e i vincoli di bilancio dei propri consumatori.

### 2.4.1 Gradi di discriminazione

Possiamo individuare, sotto il profilo economico, tre diverse modalità attraverso cui le imprese riescono a raccogliere le informazioni rilevanti per la determinazione del prezzo di riserva:

- *«discriminazione di primo grado»* o «perfetta discriminazione»: è il caso in cui l'impresa riesce a fare a ciascuno dei suoi acquirenti un prezzo differente, come avviene nel caso dei rivenditori di automobili che offrono uno stesso modello a prezzi diversi al termine di trattative individualizzate. Il prezzo fissato sarà corrispondente alla stima del prezzo massimo che il consumatore è disposto a pagare.

  Si tratta di un caso che in economia è stato considerato come teorico, in considerazione dell'elevato set informativo di cui le imprese dovrebbero disporre

per attuare una simile strategia. L'enorme disponibilità di dati individuali connessa all'avvento dei big data sta rendendo sempre più concreta la possibilità per gli operatori online di attuare strategie di perfetta discriminazione di prezzo.

"Storicamente, la discriminazione dei prezzi di primo grado è stata molto difficile da attuare, principalmente per motivi logistici. Con i progressi della tecnologia e la raccolta di big data, potrebbe essere che diventerà più facile da fare. " [John Gourville, professore di economia aziendale alla Harvard Business School, citato da Forbes (14 aprile 2014)].

Come detto, l'avvento di informazioni ricche e molto dettagliate sulla singola navigazione web ha iniziato a rendere possibili forme sofisticate di prezzi personalizzati, finora considerate troppo impegnative dal punto di vista informativo per essere implementate.

Le analisi condotte sui big data consentono quindi alle imprese di prevedere il comportamento economico e sociale di ciascun cliente e ciò rende sempre più praticate le strategie di differenziazione dei prezzi dei beni e servizi venduti online a seconda delle caratteristiche dei consumatori.

- *«discriminazione di secondo grado»* o del «menu pricing»: è il caso in cui l'impresa fissa i prezzi da praticare a seconda delle caratteristiche associabili al suo bene, per poi aspettare che i consumatori rivelino i loro prezzi di riserva proprio compiendo una scelta rispetto a quelle caratteristiche. Le imprese sono, infatti, in grado di offrire moltissime offerte rispetto alle quali i consumatori si auto-selezionano. (es: sconti quantità, i prezzi praticati dalle compagnie aeree che prevedono prezzi ridotti per chi acquista i biglietti con largo anticipo).

- «discriminazione di terzo grado» o di «group pricing»: si tratta di strategie utilizzate quando è possibile fissare prezzi differenti, in base a caratteristiche facilmente osservabili del consumatore, come l'età, il genere e talvolta l'occupazione (es: prezzi scontati per fasce di consumatori più giovani o mature, ingressi gratuiti a seconda del genere, biglietti scontati per categorie come gli studenti).

### 2.4.2 Impatto prezzi personalizzati su welfare

In linea generale, le strategie di discriminazione del prezzo sono considerate come un fenomeno in grado di aumentare, o comunque non ridurre, il benessere sociale e, quindi, per certi versi sono state in passato giudicate auspicabili.

Attraverso tali pratiche, infatti, da un lato si può aumentare il benessere sociale perché è possibile realizzare la vendita di beni che altrimenti non sarebbero mai stati venduti in assenza dell'individuazione, da parte del venditore, di consumatori disposti a pagare un prezzo più alto rispetto ad altri; dall'altro lato, si riesce a garantire una maggiore partecipazione allo scambio tramite il coinvolgimento anche di quei consumatori che, presentando una bassa disponibilità a pagare, senza discriminazione non avrebbero partecipato allo scambio dal momento che avrebbero dovuto sostenere un prezzo effettivo superiore rispetto a quello massimo che sono disposti a pagare.

La pratica della discriminazione di prezzo, tuttavia, è stata, in alcuni casi, criticata, facendo leva su una questione di equità piuttosto che su basi puramente economiche. In maniera intuitiva, infatti, offrire un medesimo bene o servizio a prezzi diversi causa un danno a coloro che presentano una disponibilità a pagare superiore. Se la discriminazione è realizzata in modo efficace, inoltre, può permettere all'azienda di raggiungere il massimo profitto disponibile sul mercato, lasciando i consumatori senza nessun beneficio dalle transazioni.

La discriminazione di prezzo in concorrenza non rappresenta una minaccia per l'efficienza dei mercati e la distribuzione della ricchezza, ma diventa problematica quando praticata da monopolisti e imprese in posizione dominante.

Tuttavia, la discriminazione perfetta del monopolista costituisce comunque una opzione preferibile rispetto al divieto di discriminazione e dunque al prezzo di monopolio.

In particolare, infatti, la discriminazione di 1° grado, idealmente, permette di fissare tanti prezzi quanti sono i consumatori. Fissando tanti prezzi, l'impresa monopolista si appropria di tutto il surplus del consumatore che diventa nullo e si trasforma interamente in profitto.

Il benessere collettivo, il cosiddetto welfare, dovrebbe rimanere invariato, ma se si produce un aumento dell'output (e quindi del numero di consumatori che riescono a godere del bene) il welfare aumenterebbe.

Seguendo, quindi, questi ragionamenti, anche la discriminazione perfetta delle imprese in posizione dominante è preferibile a uno scenario di divieto di discriminazione, qualora si assicuri che tale discriminazione perfetta abbia prodotto un aumento dell'output.

$$p^{\circ}\Delta x - c\Delta x \geq \Delta W \geq p^{1}\Delta x - c\Delta x$$

$$(p^{\circ} - c)\sum \Delta x_i \geq \Delta W \geq (p^{1} - c)\sum \Delta x_i$$

$$\Delta W \geq 0 \qquad se \qquad \sum \Delta x_i \geq 0$$

## 2.5 I big data e la pubblicità personalizzata

La pubblicità online è forse l'esempio più familiare di come le imprese utilizzano i ricchi dati forniti dagli utenti.

Le novità della raccolta di dati online sono la portata dei dati generati e ottenuti, la precisione con cui l'impresa può collegare un'azione a un consumatore specifico e l'enorme quantità di informazioni.

Facebook e, in generale, i social network vivono dei nostri like, utili a profilare i nostri gusti e le nostre preferenze, insieme alle informazioni comunicate dai nostri smartwatch, i film che più ci piacciono scelti su Netflix, i cookie memorizzati sui nostri pc che svelano i siti di nostro gradimento.

Tutte queste informazioni sono in grado di disegnare una precisissima mappa dei nostri gusti, delle inclinazioni al consumo, grazie alla quale sarà possibile, per le tech companies, inviarci pubblicità personalizzata e offerte su misura che incontreranno il nostro apprezzamento, spingendoci all'acquisto.

Prima dell'acquisto online, i negozi raramente osservavano carrelli della spesa abbandonati, dichiarazioni delle preferenze dei clienti o un elenco completo di tutti gli acquisti passati.

Le due principali differenze tra online e offline advertising sono: targetability e measurability. Questi due aspetti principali hanno contribuito a rendere le piattaforme, come Google e Facebook, tra le società in più rapida crescita e più innovative nell'economia degli Stati Uniti.

### 2.5.1 Targetability

La collezione e l'uso dei dati permette di determinare quali consumatori potrebbero essere maggiormente influenzati da una particolare pubblicità. I dati sono utilizzati per targetizzare gli utenti, mostrandoli pubblicità personalizzate ed aumentando la probabilità che l'advertising abbia successo. Un esempio di targeting è quello effettuato dalla televisione che predilige alcuni contenuti per determinate fasce orarie rispetto ad altri, basandosi sulla tipologia di pubblico.

L'online advertising rappresenta un'evoluzione più raffinata di questa realtà. Il targeting non è più legato alla fascia oraria, ma al tracciamento degli utenti che utilizzano la piattaforma; infatti, su internet è molto più facile mostrare annunci diversi a diversi sottogruppi poiché le pubblicità sono inviate ai singoli computer. Esistono, oggi, varie tecnologie che consentono alle aziende di identificare e tracciare i singoli clienti e, in tal modo, offrire loro prezzi personalizzati.

Amazon ha trovato modi innovativi alternativi di fissare prezzi personalizzati, attraverso l'utilizzo del "Gold Box". A ciascun consumatore è fornito l'accesso a una scatola d'oro ben visibile con il suo nome sulle pagine web di Amazon. L'apertura della Gold Box consente di visualizzare un numero limitato di prodotti con sconti speciali, non disponibili al di fuori della Gold Box. Poiché gli articoli offerti nella Gold Box sono diversi per i diversi consumatori, ciò consente ad Amazon di addebitare prezzi personalizzati.

La capacità delle imprese di acquisire informazioni sui gusti dei consumatori sta rapidamente avanzando. Quasi tutti i principali rivenditori, dalle catene di alimentari alle banche di investimento al servizio postale degli Stati Uniti, hanno ora un dipartimento di "analisi predittiva", riservato non solo allo studio delle abitudini di acquisto dei consumatori, ma anche delle loro abitudini personali.

Le imprese tengono traccia del comportamento di acquisto e navigazione web dei consumatori, grazie all'utilizzo diffuso di "cookie di viaggio", basati sugli accessi dei consumatori su siti popolari, ad esempio siti di compagnie aeree, Facebook o Amazon.

Una volta effettuato l'accesso, il cookie segue il cliente ovunque si trovi sul Web.

BlueKai, una startup ben nota per la raccolta e l'ordinamento dei dati sui consumatori, è una delle aziende più note che forniscono il software che consente ai siti Web di tracciare i propri utenti in modo da assegnarli a specifici segmenti di mercato.

### 2.5.2 Measurability

Oltre al targeting, gli inserzionisti online raccolgono e analizzano i dati per misurare l'efficacia dell'annuncio. Questa pratica funziona per due motivi. In primo luogo, la piattaforma online consente a un'impresa di collegare la visione di un consumatore di un annuncio pubblicitario, la sua reazione (click o non click) ed il suo comportamento successivo (acquisto o non acquisto). In secondo luogo, la piattaforma online facilita gli esperimenti sul campo in cui le aziende mostrano casualmente ai diversi consumatori pagine Web diverse.

# 3. CONCENTRAZIONE DEI MERCATI

I mercati digitali sono spesso caratterizzati da dinamiche winner take all, dove il vincitore si prende tutto il mercato. Vi è, quindi, un'elevata concentrazione di mercato, una dispersione dei profitti maggiore e forti divari economici tra le imprese.

In questo capitolo, analizzeremo tutte le possibili fonti di una maggiore concentrazione nel mercato digitale.

Il fatto che i servizi digitali siano spesso offerti gratuitamente non significa che le preoccupazioni sulla concentrazione non siano fondate.

## 3.1 Switching cost come fonte di vantaggio per l'operatore storico?

A volte la concorrenza nei mercati è limitata dalle difficoltà degli utenti di cambiare operatore o di utilizzare un servizio differente. Può esserci, in alcuni casi, addirittura una strategia aziendale atta a impedire che gli utenti possano trasferire i propri dati ad un concorrente per non perdere la clientela. In questi casi le imprese, alzando i cosiddetti switching cost, ostacolano la concorrenza ed impediscono l'ingresso di nuovi operatori. Tali costi sono tanto più elevati quanto più è difficile per un consumatore migrare da un fornitore di servizio a un altro.

Switching cost molto alti possono provocare un effetto di lock-in, che consiste nel fatto che gli utenti, "bloccati" da un determinato operatore, sono costretti ad usare i suoi servizi indipendentemente dalla potenziale maggiore qualità offerta da un gestore differente. Soprattutto per i servizi che utilizzano i dati degli utenti in maniera rilevante, può risultare piuttosto agevole per le imprese limitare la capacità di cambiare fornitore, bloccando la migrazione dei dati alla concorrenza.

I consumatori che adottano tecnologie precedenti sono "bloccati" nel caso in cui una nuova tecnologia, anche superiore, si presenti in una fase successiva. Il cambio di costi e, più in generale, i consumatori "incagliati" sono una chiara fonte di vantaggio per l'operatore storico.

Tuttavia, mentre gli switching cost hanno chiaramente giocato un ruolo importante nello sviluppo iniziale di software e hardware, i mercati digitali di oggi sono generalmente caratterizzati da switching cost molto bassi.

Iscriversi a un servizio, installare un'app o registrarsi su un sito Web non richiede investimenti in nuove apparecchiature e quindi ciò che ne consegue è che gli switching cost, essendo molto bassi, non sono la causa principale di una concentrazione nei mercati digitali.

## 3.2 I big data e il potere di mercato

In rapporto agli utilizzi dei dati generati e raccolti da un'impresa, si delineano due diversi scenari, non mutualmente escludenti.

In primo luogo, è possibile che le imprese generino o raccolgano dati digitali per rivenderli in tutto o in parte. In questo caso, bisognerà controllare se, nei successivi mercati a valle della commercializzazione dei dati digitali, una o più imprese riescano ad acquisire potere di mercato. Tuttavia, quest'ultima analisi rappresenterà un'indagine distinta da quella relativa all'accertamento dell'esistenza di un certo potere di mercato nel mercato della generazione e raccolta dei dati digitali. In altri termini, di volta in volta, si dovrà definire i confini dei mercati per la commercializzazione dei dati, così da verificare come essi si collochino e comportino rispetto al mercato o ai mercati per la generazione o raccolta dei dati.

In secondo luogo, è possibile che le imprese scelgano di usare internamente i dati da loro generati e raccolti per poi ideare o perfezionare nuovi beni e servizi. In questo caso, bisognerà procedere con l'individuazione dei mercati di questi beni e servizi, mercati in cui le imprese potranno anche avere un notevole potere di mercato, ma non perché depositarie di grandissimi volumi di dati, o perché uniche detentrici di una particolare categoria di dati, ma in virtù di quanto i loro beni e servizi superano quelli dei rivali.

Di conseguenza, non sembra ragionevole ritenere che i big data in quanto tali possano da soli attribuire a un'azienda il potere di agire indipendentemente dai suoi concorrenti nei mercati dei suoi beni e servizi. Il fatto che un'impresa disponga di big data non implica necessariamente che quell'impresa disponga di potere di mercato nel mercato o nei mercati della generazione e raccolta di quei dati. Tuttavia, è ovvio, anche per quanto detto nei capitoli precedenti, che il possesso dei big data è ciò che permette a un'impresa di migliorare i suoi beni e servizi al punto di sbaragliare la concorrenza.

## 3.3 Big data come fonte di vantaggio competitivo

Si discute delle barriere alla raccolta dei dati, oscillando tra due posizioni estreme.

Per un verso, vi è chi ricorda (Thierer, Lerner, Lambrecht, Tucker,...) come la non-rivalità dei dati li renda facili da ottenere, a costi di produzione e distribuzione molto bassi. Gli utenti di internet e degli smartphone lasciano, infatti, le proprie "tracce digitali" in molte delle piattaforme con le quali contemporaneamente interagiscono (fenomeno del multi-homing) e quindi è molto semplice raccogliere dati digitali. Di conseguenza, questa parte della letteratura sottolinea come i dati non posseggano un grande valore e come, anche quando raccolti in grandi volumi, essi siano facilmente reperibili.

Nel dettaglio, poi, questa dottrina sostiene anche che un'impresa, che decide di entrare in un mercato nel quale è già presente un incumbent ricco di dati, potrebbe in breve tempo competere ad armi pari, perché, superato un certo ammontare di dati, i dataset si caratterizzano per economie di scala decrescenti, per cui ogni unità di dato aggiunto da un certo punto in poi riduce il suo valore fino a non aggiungerne più. Tutto ciò renderebbe meno efficienti le economie di scopo e più in generale non razionale dal punto di vista del titolare del dato, in posizione dominante, utilizzare il proprio potere di mercato per attuare condotte unilaterali (abusi) per ridurre o falsare la concorrenza.

Attraverso lo studio di molti casi empirici, si è negata, quindi, l'esistenza di un first mover advantage, esaltando il fenomeno del multi-homing, ricordando come anche l'*autenticità* e l'*utilità* di un dato, inoltre, siano difficili da verificare a priori.

Così, questa dottrina è arrivata a negare l'ipotesi secondo la quale i mercati digitali si concentrerebbero, portando ad esempio casi di imprese che oggi hanno acquisito importanti quote di mercato senza essere state le prime ad operare nei relativi mercati digitali.

D'altra parte, questa dottrina, seppur correlata dallo studio di molti casi empirici, non sembra rispecchiare ciò che invece avviene in realtà. I mercati digitali sono concentrati. Bisogna solo capire il perché. Secondo una diversa corrente di pensiero (Newman, Wright, Kaiser, Parker, …), vi è chi sostiene come, per lo meno con riguardo all'operatività delle piattaforme digitali, leader spesso indiscussi nel mercato dove si raccolgono i Big Data (mercato a monte), i dati siano diventati una risorsa imprescindibile, ossia una risorsa senza la quale i modelli di business di queste piattaforme non potrebbero funzionare.

Proseguendo lungo questa linea, questi autori hanno evidenziato come almeno alcuni tipi ed insiemi di dati raccolti dalle piattaforme digitali possano diventare risorse inimitabili e comunque inaccessibili a molti.

Inoltre, su alcuni utilizzi specifici dei dati come nell'analisi predittiva basata sulle probabilità derivante da un numero maggiore di osservazioni, più alta è la misura della probabilità e quindi dei dati disponibili, tanto più è affidabile la previsione: una sorta di "learning by doing" a tutto vantaggio dell'utilizzatore, che spiegherebbe ad esempio, nel caso di Google, la superiorità del suo motore di ricerca. Il concetto di learning by doing va sicuramente a scontrarsi con l'idea di dataset con economie di scala decrescenti della precedente dottrina. È forse questo il punto cruciale su cui si scontrano maggiormente le due dottrine e, rispetto a questi due poli del dibattito, la dottrina più sofisticata ha compreso la necessità di procedere caso per caso, distinguendo le ipotesi nelle quali i dati sono disponibili e liberamente appropriabili da quelle in cui i dati raccolti da alcune imprese sono non duplicabili e comunque riconoscendo come la raccolta dei dati digitali richieda infrastrutture, algoritmi e competenze costose.

## 3.4 Big Data nei mercati a valle

Al pari di qualsiasi altra impresa interessata ad approvvigionarsi di input, le imprese determinate a raccogliere dei dati digitali possono o decidere di integrarsi verticalmente, entrando nei mercati della generazione e raccolta dei dati, cercando di competere quindi con le grandi piattaforme digitali, o scegliere di acquisire quei dati dal mercato, proprio da quest'ultime.

Quest'ultima strategia comporta necessariamente un costo, che gli esperti di diritto antitrust sono soliti definire "barriere all'ingresso", perché questi costi lavorano come ostacoli che i potenziali entranti devono superare per iniziare a operare in un determinato mercato.

I big data, infatti, risulterebbero come una barriera posta a difesa delle imprese già presenti nel mercato a valle. Al crescere dei dati disponibili a un'impresa, risultano maggiori le economie di scala e le economie di scopo di cui quell'impresa beneficia nella produzione e commercializzazione dei suoi beni e servizi.

Si può, quindi, ipotizzare che i big data riducano i costi delle imprese già presenti nel mercato andando così a rinforzare il loro potere di mercato e a rendere più probabile l'eventualità che il mercato rilevante risulti altamente concentrato.

## 3.5 Big data nei mercati a monte

Buona parte del vantaggio competitivo delle piattaforme digitali deriva dalla collezione e dall'uso dei dati.

I big data sono assimilabili a delle risorse ritenute essenziali per la realizzazione di un determinato prodotto o servizio. Ciò rende l'impresa o piattaforma che detiene quei particolari dati nella condizione privilegiata di godere di un vantaggio competitivo assoluto che rende la sua posizione di mercato meno contendibile sia all'interno del mercato della raccolta dei dati che all'interno dell'intera filiera logistica.

Grazie ai dati raccolti dalle piattaforme, queste ultime si creano una conoscenza più approfondita della domanda e, dunque, non solo riescono a raffinare i propri servizi in relazione alle caratteristiche dei singoli consumatori, ma possono anche offrire agli inserzionisti spazi pubblicitari sempre più idonei a catturare utenti con preferenze e abitudini di acquisto compatibili con i prodotti e servizi che essi intendono fornire. Viene così a determinarsi un doppio effetto di feedback positivo a favore delle piattaforme che, da un lato, perfezionando i propri servizi, aumentano i loro utenti, i quali a loro volta mettono a disposizione delle piattaforme sempre più dati e, dall'altro lato, migliorando la loro offerta di pubblicità comportamentale, aumentano i loro inserzionisti, i quali dal loro canto sono disposti a pagare sempre di più perché sono messi in contatto proprio con i consumatori maggiormente propensi ad acquisire i loro beni.

## 3.6 Aspetto comportamentale: effetti di rete

Un tema comune negli ultimi lavori è quello di guardare al "coordinamento" dei consumatori.

Già nel capitolo precedente avevamo affrontato il tema delle esternalità di rete all'interno delle multi-sided platforms. Gli effetti di rete sono un'ovvia fonte di concentrazione, perché un numero maggiore di utenti migliora l'attrattiva dell'impresa dominante, portando a un numero ancora maggiore di utenti. In casi estremi, gli effetti di rete provocano il monopolio dei mercati.

Prendiamo come esempio un concorrente con una tecnologia leggermente superiore che vuole sfidare un operatore storico. Il concorrente è una filiale di una grande azienda con una reputazione consolidata e i servizi offerti sono molto simili a quelli dell'operatore storico. Non ci sono switching cost e il prezzo offerto dal concorrente non è superiore a quello offerto dall'operatore storico.

Questa situazione sembra catturare meglio il panorama attuale con diverse aziende che non riescono a spostare gli operatori storici nonostante tasche profonde, reputazione consolidata e marchi famosi. Ne sono esempi gli sforzi di Microsoft per sfidare Google nel mercato dei motori di ricerca o lo sforzo di Google per sfidare Microsoft nelle app per la produttività in ufficio.

Un operatore storico è detto "focale" se i consumatori fanno le loro scelte ipotizzando che tutti gli altri consumatori non cambieranno. Si può facilmente dimostrare che in un modello statico (one-shot) tra un operatore "focale" relativamente inefficiente e un concorrente "non focale", l'operatore storico conquista il mercato "troppo spesso".

Il motivo è che l'unico modo per il concorrente di guadagnare quote di mercato è quello di renderlo una strategia dominante per la migrazione dei consumatori, ovvero i consumatori dovrebbero trovare conveniente migrare, qualunque cosa facciano gli altri consumatori. Pertanto, l'unico modo in cui un'impresa entrante può guadagnare quote di mercato è offrire servizi stand-alone aggiuntivi, indipendenti dalla rete, o applicare prezzi relativamente bassi o, se possibile, potenzialmente negativi.

Chiaramente, il concorrente sarebbe disposto a investire risorse per indurre i consumatori a cambiare solo se si aspetta di recuperare tali investimenti in futuro. Ciò porta all'inefficienza statica per cui alcuni concorrenti di qualità superiore potrebbero non riuscire a conquistare il mercato.

Halaburda, Jullien e Yehezkel (2018) e Biglaiser e Cremer (2018) si sono occupati di questo problema, includendo considerazioni anche dinamiche, osservando la concorrenza tra un operatore storico e una sequenza di potenziali concorrenti per un gran numero di periodi.

Le imprese concorrenti potrebbero avere un incentivo a investire, incorrendo in perdite nel breve periodo, per raccogliere profitti in futuro. Persuadere i consumatori di fronte a convinzioni sfavorevoli è costoso, ma d'altra parte potrebbe consentire in seguito di recuperare quegli investimenti attraverso i prezzi di monopolio. Rispetto al modello statico, i partecipanti dovrebbero essere disposti ad incorrere in perdite maggiori nel breve periodo per conquistare il mercato.

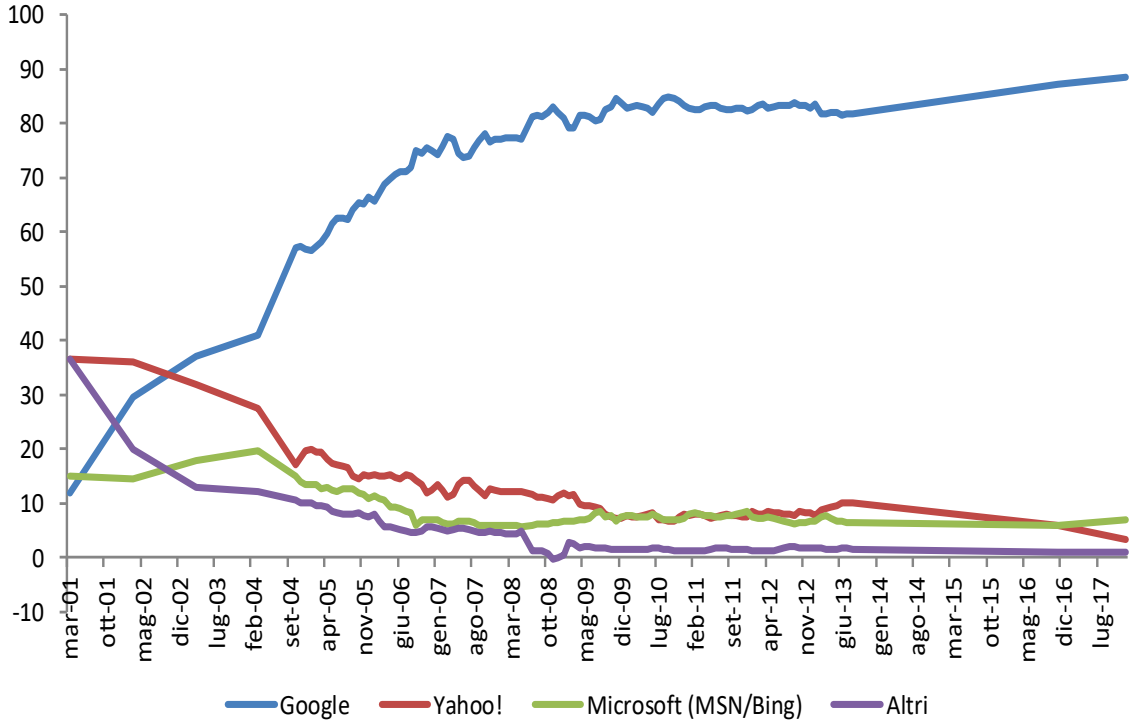
Tuttavia, l'inefficienza statica si estende anche all'impostazione dinamica: gli operatori storici focali e di qualità inferiore possono spesso rimanere dominanti. Pertanto, le considerazioni a lungo termine non ripristinano necessariamente risultati efficienti.

In particolare, Halaburda, Jullien e Yehezkel (2018) mostrano che questo risultato inefficiente si verifica quando le aziende non sono molto pazienti.

Le esternalità di rete hanno un effetto non solo sul valore della rete stessa, ma anche sulla profittabilità delle imprese. Infatti, mentre i costi tendono a crescere al più linearmente, i potenziali ricavi aumentano esponenzialmente.

Nel 2019, le prime cinque piattaforme online (Apple, Amazon, Google, Microsoft e Facebook) hanno fatturato oltre 700 miliardi di dollari, con un utile netto pari a circa 110 miliardi, quindi superiore al 15% dei ricavi.

L'andamento storico della quota di Google nel mercato dei motori di ricerca **(Figura 11)** mostra in modo abbastanza eloquente la dinamica winner take all che caratterizza i mercati digitali.




*"Figura 11: quote di mercato nel mercato dei motori di ricerca"*

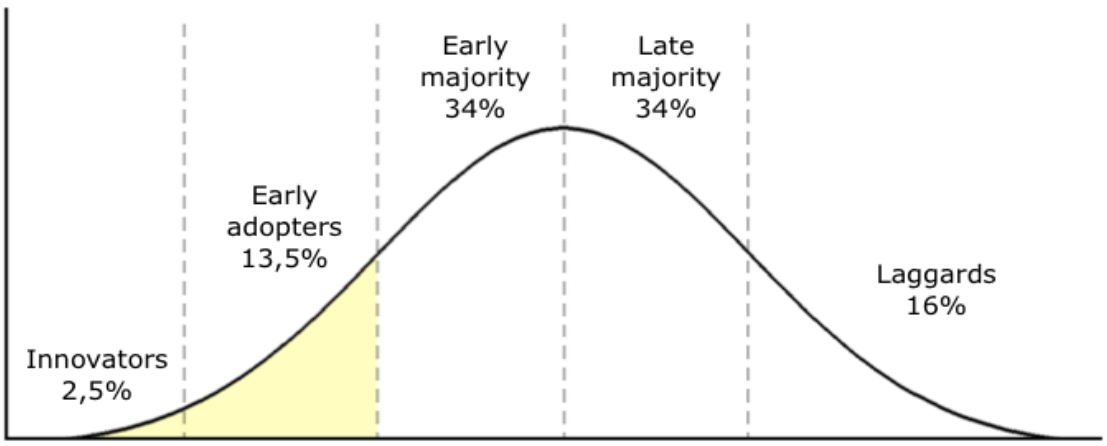
Biglaiser, Cremer e Veiga (2018), inoltre, hanno identificato, come fonte di vantaggio per l'operatore storico, gli incentivi dei consumatori ad aspettare fino a quando la base di utenti del concorrente è abbastanza grande. Nel loro modello, i consumatori non scelgono contemporaneamente se migrare al concorrente, ma, in ogni momento, arriva l'opportunità di migrare con qualche probabilità positiva.

Questo modello di equilibrio cattura un fatto ben noto nel settore, che è riluttanza dei consumatori a unirsi presto alle piattaforme concorrenti.

Come spiega, infatti, il modello di Rogers, i consumatori sono divisi in categorie:

- Innovators (2,5%) = coloro che credono fin dall'inizio nell'innovazione e che sono disposti a pagare molto;
- early adopters (13,5%) = coloro che acquistano quando sono presenti i primi contenuti della tecnologia;
- early majority (34%) = coloro che acquistano quando la tecnologia è già più facile da utilizzare;
- late majority (34%) = coloro che acquistano quando la tecnologia diventa uno standard;
- laggards (16%) =adottano la tecnologia proprio se non c'è alternativa.

Essere un "early adopter" comporta il costo di rinunciare agli effetti di rete collegati alla base di utenti più ampia. Si ritiene che, anche laddove dei consumatori volessero cambiare fornitore di beni e servizi, non sarebbero incentivati a farlo perché, così facendo, potrebbero passare a un fornitore che, privo di quei dati, risulterebbe meno capace di proporre un'offerta ritagliata sulle loro preferenze.
Il modello di Rogers è molto utile per farci un'idea qualitativa del grado di adozione di quella particolare piattaforma.




*"Figura 12: modello di Rogers"*

## 3.7 Ignoranza del consumatore e asimmetria informativa

Nei mercati per la generazione e raccolta dei big data, le imprese godono di una esternalità positiva radicata nell'ignoranza dei consumatori che determina un eccessivo trasferimento di dati dai consumatori verso le imprese. I consumatori non sono consapevoli di molti dei meccanismi che governano l'economia digitale, come il valore dei dati in generale. Essi, infatti, non sanno calcolare il livello di dati che dovrebbero mettere a disposizione delle imprese per massimizzare la propria utilità e il mercato tende così a fallire perché raggiunge l'equilibrio in corrispondenza di un livello sub-ottimale di dati, perché eccessivo.

La maggior parte dei consumatori a cui vengono offerti dei prezzi personalizzati non sa né del processo data driven e data based attraverso il quale quei prezzi sono stati fissati, né della loro natura di "prezzi diversi per ogni compratore". Una forma di duplice ignoranza, quest'ultima, che suscita qualche riflessione sulla natura eventualmente scorretta della pratica dei prezzi personalizzati.

I big data limitano il funzionamento del mercato contribuendo ad incrementare le asimmetrie informative che separano imprese e consumatori. Grazie a i big data, infatti, le imprese sono nelle condizioni di accedere a molte più informazioni rispetto a quelle disponibili ai consumatori: l'elaborazione dei dati consente loro di comprendere preferenze di cui gli stessi consumatori possono non essere consapevoli.

Complessivamente, dunque, queste forme di ignoranza/inconsapevolezza dei consumatori impediscono che la pressione competitiva operi da pungolo sia nei mercati della generazione e raccolta dei dati, sia nei mercati della commercializzazione dei beni e servizi realizzati dalle imprese che controllano quei dati.

È possibile ipotizzare interventi tesi a mitigare le asimmetrie informative attraverso una puntuale definizione degli adempimenti richiesti per l'uso dei dati personali degli utenti, ma tale approccio, negli anni, si è dimostrato non particolarmente efficace. Gli interventi, con i quali è aumentato l'obbligo di informare per le imprese, hanno finito per tradursi in adempimenti prevalentemente formali, inadeguati a fronteggiare il tema sostanziale di fondo dell'adeguata informazione del consumatore e delle scelte consapevoli compiute dagli stessi.

Tuttavia, aumentare la consapevolezza del consumatore sulle modalità in cui sono raccolti e analizzati i dati potrebbe aiutare le aziende a creare un rapporto di fiducia con i clienti,

mitigando così ex-ante eventuali problemi di privacy. L'educazione del consumatore potrebbe, quindi essere considerata come un modo per fidelizzare il cliente.

Sebbene implichi costi anche elevati, cercare per lo meno di ridurre le asimmetrie informative dovrebbe essere, in ogni caso, visto come uno dei principali obiettivi della regolazione. È per questo che sono state, recentemente, avviate alcune iniziative educative finalizzate a far conoscere maggiormente i mercati dei dati digitali e a diffondere la consapevolezza dell'importanza di un uso responsabile di internet. Programmi educativi possono, quindi, talvolta aiutare anche se devono essere usati con cautela e valutando sempre costi e benefici associati.

## 3.8 I big data tra monopoli naturali?

I comportamenti delle imprese non sono gli unici fattori che possono modificare il funzionamento del mercato. Può anche accadere che caratteristiche dei processi produttivi e degli agenti economici coinvolti nelle transazioni impediscano il pieno svolgersi del gioco concorrenziale. È in presenza di questi ultimi fallimenti di mercato che la regolazione trova la sua giustificazione economica, candidandosi a svolgere un ruolo fondamentale.

Il fallimento di mercato per eccellenza è la presenza di un monopolio naturale, che ricorre ogni volta che la struttura dei costi rende razionalmente conveniente l'esistenza di un'unica impresa incaricata della produzione. Il monopolio, in questi casi, è ottimale, come nei settori delle telecomunicazioni e la regolazione diventa necessaria per garantire che i consumatori non siano danneggiati da prezzi troppo elevati.

Sebbene le tecnologie dell'informazione possano comportare forme di concorrenza imperfetta, i costi delle attività di raccolta, generazione e analisi dei dati digitali possono essere minimizzati anche quando le imprese attive nel mercato sono più di una e quindi ciò ci porterebbe a rifiutare l'ipotesi che i mercati digitali con i big data, possano essere associati a dei monopoli naturali.

# 4 Il mercato delle APP

Uno dei principali meccanismi attraverso cui i consumatori cedono dati digitali avviene tramite il *download* e il successivo utilizzo delle app, cioè di quei programmi, o pacchetti di programmi che ciascun utente installa, o trova già installati, sul proprio *device* per lo svolgimento di specifiche attività.

Nel giro di pochi anni, l'ecosistema collegato alle APP ha mostrato una crescita esponenziale. Il fenomeno è abbastanza recente, dato che il primo APP *store*, creato dalla società Apple, è stato lanciato nel 2008, un anno dopo l'uscita del primo modello di *iPhone*. Dopo solo tre anni dalla nascita del primo APP *store*, erano già disponibili per gli utenti circa un milione di app.

Secondo i calcoli della società di analisi *App Annie*, l'ecosistema delle APP (comprensivo di applicazioni a pagamento, pubblicità e acquisti mobile) valeva oltre 1.300 miliardi di dollari nel 2016, mentre per il 2021, secondo le stime, dovrebbe superare oltre 6.300 miliardi di dollari, con una crescita del 385%. Tale previsione è guidata principalmente da due forze: la diffusione di *device* mobili ma, soprattutto, la crescita del tempo medio speso nell'uso di APP, in media 3 ore al giorno per utente a fine 2017.

## 4.1 Il ruolo degli app store

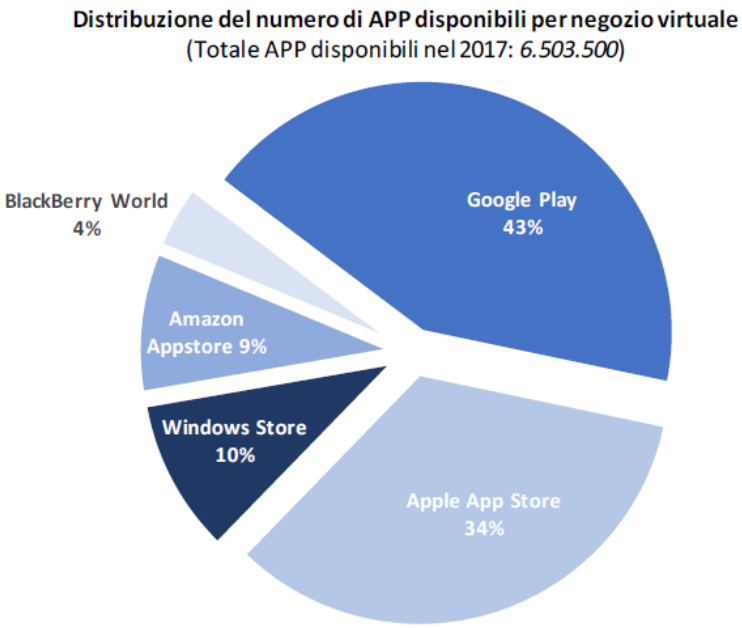
La maggioranza delle applicazioni si trova in veri e propri negozi virtuali denominati APP store: si tratta di piattaforme di distribuzione attraverso cui sono rese disponibili agli utenti le APP.

Per motivi legati ai fenomeni di integrazione verticale, i negozi virtuali possono presentare delle incompatibilità con i sistemi operativi: ad esempio, l'APP *store* di *Google* funziona su *Android* ma non su terminali *iOS* e, viceversa, il negozio di *Apple* opera sul proprio sistema operativo ma non su *Android*. Tale aspetto è molto rilevante perché produce effetti sul mercato e sui relativi assetti concorrenziali: si rafforza la posizione di mercato dei principali *player* integrati verticalmente (in particolare *Google* e Apple). Inoltre, la posizione detenuta da questi ultimi sia nei sistemi operativi sia negli APP *store* rappresenta, infatti, una situazione di enorme vantaggio nel mercato dei *big data*, in particolare nella fase della loro raccolta.

La varietà e la disponibilità di APP in uno *store* sono cresciute in maniera esponenziale dal 2008 ad oggi: nel 2017, sull'APP *store* di Apple erano disponibili 2,2 milioni di app, mentre su *Google Play* 2,8 milioni. Un numero nettamente inferiore di APP è presente negli altri negozi virtuali.

Se, dal punto di vista dei ricavi, l'*Apple store* genera un fatturato ancora superiore a *Google Play*, dal punto di vista dei volumi la situazione è esattamente opposta (**Figura 13**). Infatti, *Apple* ha adottato un modello chiuso, con un maggior margine di guadagno, mentre *Google Play* è leader in termini di numero di APP disponibili e caricate dagli sviluppatori e di *download* effettuati dagli utenti finali. Il negozio di APP "*Google Play*" è, di fatto, quello più grande in termini di disponibilità di APP e il 94% circa di queste consente il *download* gratuito e forme di *in-APP purchase* (cioè vendite aggiuntive dopo il *download*), mentre il 70% circa risulta totalmente gratuito. Ciò ha generato ricavi complessivi per 9,8 miliardi di dollari.

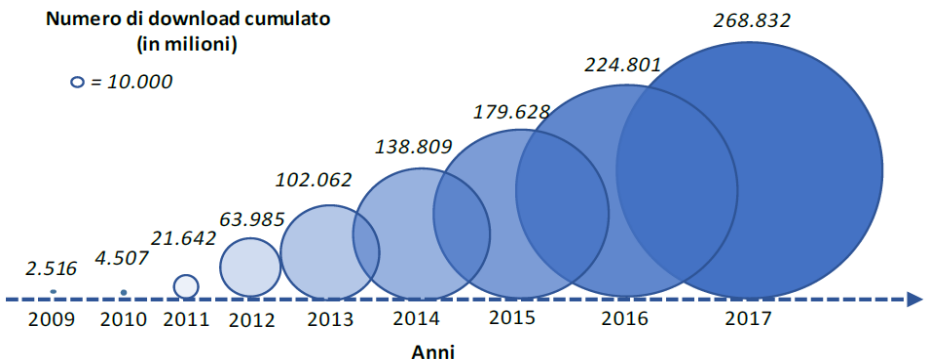
Il segmento degli APP *store*, quindi, appare caratterizzato da un elevato livello di concentrazione, con due operatori (*Google* e *Apple*) in una posizione di potere di mercato, anche in virtù dell'integrazione verticale e della quota detenuta nel collegato mercato dei sistemi operativi su *device* mobili.




*"Figura 13: quote del mercato in volume (2017)"*

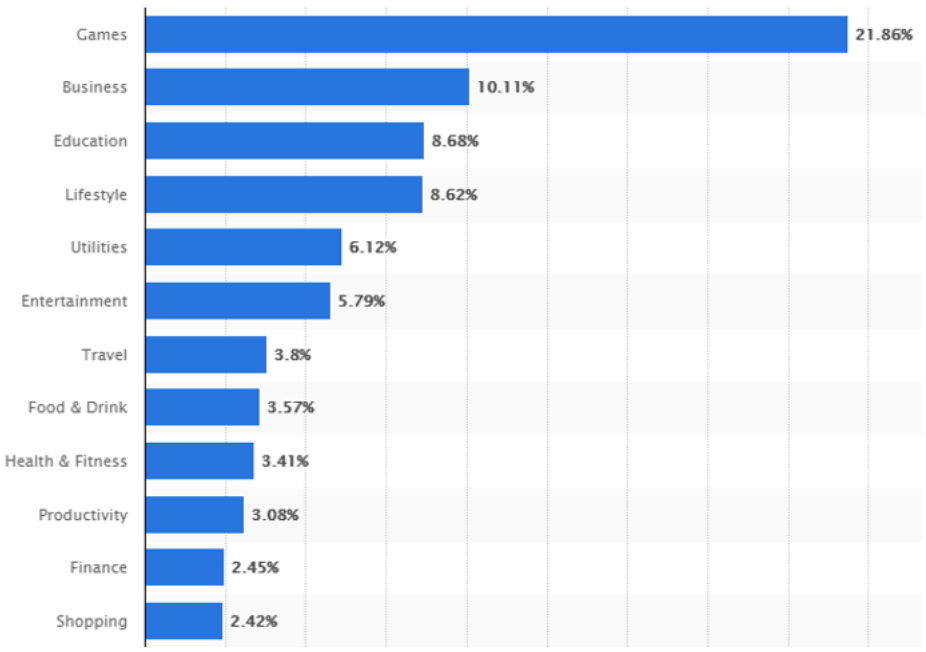
Secondo il sito specializzato nel settore, *AppAnnie*, l'utilizzatore medio di smartphone ha sul proprio *device* 80 APP e ne usa in media 40 al mese.

La **Figura 14** mostra la rapida crescita del numero di APP scaricate dagli utenti su tutti gli APP *store* esistenti: in soli 8 anni, il numero di *download* è aumentato più di 10.000 volte rispetto al valore del 2009, mentre nel 2018 si è rilevata una crescita del 19,5%.




*"Figura 14: Numero di app scaricate nel mondo dal 2009 al 2017 (in milioni)"*

La **Figura 15,** inoltre, mostra la distribuzione delle APP all'interno dello store di *Apple* per tipologia di categoria (simile distribuzione si ha anche all'interno dello store di Google Play). Ciò che emerge immediatamente è il rilevante ruolo delle app per i giochi. Quest'ultime sono le app più popolari, poiché circa il 22% di tutte le app disponibili rientra in questa categoria che, oltre a coinvolgere l'utente per più tempo, produce anche il maggior flusso di ricavi generati direttamente dalla vendita del servizio agli utenti finali.




*"Figura 15: Le categorie più popolari dell'Apple Store 2020"*

## 4.2 Il ruolo delle esternalità di rete e il mercato degli App store

L'ecosistema delle APP può essere collocato, dal punto di vista della teoria economica, nell'ambito dei mercati a due versanti (two sided market) o a più versanti (multi sided market).

Gli app store, infatti, non sono altro che delle piattaforme che comunicano con più mercati contemporaneamente e ciò che ne consegue è che si tratta, quindi, di un mercato in cui operano intense esternalità di rete.

Per il singolo utente, il beneficio derivante dall'acquisto e dall'utilizzo di un'app aumenta al crescere del numero di altri utenti che utilizzano la stessa app (**esternalità dirette**).

Il maggior beneficio per il singolo utente (utilità) aumenta la sua disponibilità a pagare, condizionando notevolmente il funzionamento del mercato, conducendolo in particolare a un maggior grado di concentrazione.

Come spiegato nel capitolo precedente, si manifestano anche le **esternalità di rete incrociate**, quando le decisioni prese dagli individui appartenenti a un lato del mercato (gli utenti) producono degli effetti sugli agenti appartenenti all'altro lato (sviluppatori) e viceversa.

Entrambi i lati del mercato hanno un beneficio dal numero di *download* effettuati tramite l'APP *store*. L'interesse degli utenti verso una particolare piattaforma, infatti, aumenta al crescere del numero di sviluppatori che vi operano e, viceversa, gli sviluppatori saranno portati a produrre APP per uno *store* all'aumentare degli utenti che vi accedono.

L'intensità delle elasticità incrociate nelle due direzioni è un fattore rilevante nelle scelte strategiche relative al prezzo delle APP fissato dalle imprese.

Circa il 94% delle app è reso, infatti, disponibile per gli utenti in maniera gratuita e, quindi, le app in questo mercato sono disposti a fissare prezzi nulli, anche a fronte ovviamente di un costo di produzione positivo. Le app, come spiegato nei capitoli precedenti, si comportano da piattaforme e, quindi, spesso, non incassano nel mercato dove sono presenti gli utenti, ma monetizzano in altri mercati come il mercato dell'advertising e il mercato della raccolta dei nostri dati.

Le esternalità di rete si manifestano principalmente attraverso il conseguimento di quella che viene definita "massa critica di utenti". Acquisita tale quantità, infatti, la rete cresce successivamente, seguendo un processo che si autoalimenta (effetto valanga o *bandwagon effect*).

Il raggiungimento della massa critica è un'ulteriore spiegazione del perché le imprese, strategicamente, in una prima fase, possono giudicare vantaggioso offrire il proprio bene o servizio ad un prezzo basso o, come spesso accade, gratuitamente. Il prezzo basso, infatti, consente di attirare più velocemente utenti e quindi di raggiungere prima la soglia critica, rispetto ad una politica di prezzo alto che in molti casi potrebbe scoraggiare l'acquisto del bene o servizio da parte dei consumatori.

Nei mercati con presenza di esternalità di rete si afferma un numero limitato di standard tecnologici, siano essi sistemi operativi, APP *store*, o singole categorie di APP e non è sempre la migliore tecnologia disponibile a prevalere sul mercato. Proprio nella fase in cui si tende a raggiungere il livello della massa critica, diventa fondamentale la spinta che i consumatori iniziali danno alla diffusione del tipo di tecnologia e, di conseguenza, la tecnologia che prevarrà è molto influenzata dalle scelte iniziali di un certo numero di utenti (effetto *lock-in*).

Tutti questi temi, tipici dei mercati a due o più versanti, possono portare alla formazione di equilibri di mercato di tipo monopolistico o di oligopoli (the winner takes all), dove il vantaggio di chi supera per primo una massa critica di utenti (non necessariamente il primo operatore a fare ingresso sul mercato) diventa sempre più importante nel tempo.

In tal senso, le APP sviluppate dai giganti tecnologici si sono evolute, rappresentando sempre più delle vere e proprie piattaforme. La APP di *Facebook*, per esempio, forte dell'enorme quantità e varietà di dati rilasciati dagli utenti, raccolti ed archiviati ad una rapidità sempre più elevata, ha introdotto nel tempo tantissime innovazioni nella fruizione di contenuti, in particolare di quelli video, e di accesso ad altri servizi in modo da far sembrare l'ambiente della sua APP sempre più simile ad una vera e propria piattaforma distributiva.

Per le caratteristiche del mercato descritte (presenza di economie ed esternalità di rete), di tale aumento hanno beneficiato soprattutto i giganti tecnologici come *Google* e *Facebook*.

## 4.3 Concentrazione nel mercato delle app

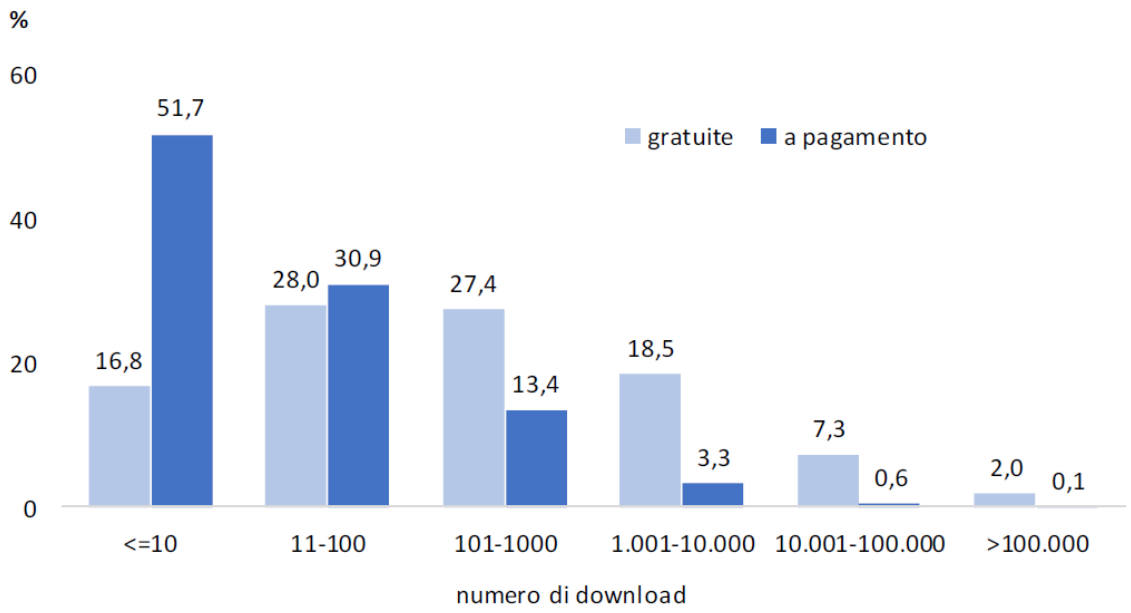
Il Servizio economico-statistico dell'Autorità, con la collaborazione del Dipartimento di Ingegneria Informatica Automatica e Gestionale dell'Università "la Sapienza" di Roma, ha sfruttato le informazioni ricavabili da un ricco *dataset* di 1.135.700 APP presenti su Google Play, circa l'80% delle APP disponibili sullo store. Il rimanente 20% fa parte di una quota residuale di app che risultano solo marginalmente diffuse, in termini di *download*, presso i consumatori finali.

Il numero medio di *download* è una buona approssimazione della domanda di APP da parte degli utenti, e, quindi, della potenziale quantità (volume) di dati raccolti dagli sviluppatori e dagli *APP store*.

La **Figura 16** mostra la distribuzione delle APP per numero di *download,* suddivise a seconda che si tratti di APP a pagamento o gratuite. Le due distribuzioni differiscono notevolmente, dal momento che le APP gratuite hanno un numero di *download* superiore. Più dell'80% delle APP a pagamento è scaricato da 1 a 100 volte, mentre, quasi un terzo delle app gratuite è scaricato più di 1.000 volte, valore che è pari a meno del 4% per quelle a pagamento.

L'andamento dei download osserva un chiaro fenomeno di "coda lunga" anche nel mercato delle APP, come per il resto del web: in effetti, senza distinguere tra APP a pagamento e non, circa il 50% delle APP è scaricato meno di 100 volte e circa il 98% meno di 100.000 volte. Ciò significa che solo il 2% di APP risulta installata da un numero considerevole di utenti. Nel mondo, solo sei APP risultano installate più di 1 miliardo di volte: Facebook, Google Gmail, YouTube, Google Maps, Google Search e Google Play Services.

Tale dato dimostra ancora una volta come, anche se ci sono tantissime app e operatori, il mercato è in realtà concentrato in poche grandi piattaforme, che raggiungono numeri di diffusione non replicabili per gli altri operatori.




*"Figura 16: distribuzione di app per numero di download"*

## 4.4 Una soluzione di mercato alle transazioni di dati: i permessi

Le informazioni che sono raccolte, potenzialmente, sono tutte quelle disponibili su un *device* mobile e riguardanti tutte le possibili attività svolte dall'utente, come il salvataggio di dati, la lista dei contatti, i video, le foto, i messaggi, la posta elettronica e le varie *password* utilizzate per accedere a specifici servizi, come quelli finanziari.

In alcuni casi, per il corretto funzionamento di un'APP è strettamente necessario consentire l'accesso ad alcune componenti hardware e software dello smartphone e ad alcune informazioni con un certo grado di sensibilità. Ad esempio, se si è interessati ad un'applicazione che riporta le previsioni meteorologiche, è necessario che essa possa accedere alle informazioni relative alla localizzazione dell'utente, per permettere all' APP di fornire le informazioni meteo sul luogo in cui si trova in quel momento l'utente.

Gli APP store si sono dotati di specifiche politiche di trattamento dei dati, sia per rispettare le norme sulla protezione dei dati vigenti nei singoli Stati, sia per creare un rapporto di maggior fiducia con i propri utenti.

Le modalità di trattamento dei dati sono rese note agli utenti prima del download di un'APP, tramite un accordo di licenza.

Il meccanismo adottato per disciplinare la parte dei rapporti che intercorrono tra utenti e sviluppatori di APP relativa al flusso di dati e al loro trattamento si basa su un sistema di permessi. Tale sistema è definito e monitorato dal gestore dello *store*.

Quando si effettua il *download* dell'APP, il sistema operativo chiede all'utente di selezionare la casella "accetto", dopo aver indicato tutte le informazioni a cui l'APP dovrà accedere. Inoltre, prima che l'APP sia installata, gli utenti possono verificare i permessi richiesti accedendo all'opzione "*dettagli autorizzazione*". La lista dei permessi viene aggiornata ogni volta che l'APP stessa viene aggiornata.
I permessi possono consentire all'APP di accedere a numerosi dati degli utenti, relativi, ad esempio, agli spostamenti, ai luoghi visitati, alle abitudini di acquisto, alle foto e ai video scattati e condivisi.

La politica dei permessi è strategicamente rilevante per il successo di un'APP, perché gli utenti si mostrano attenti alla quantità e al tipo di informazioni che sono cedute nel momento in cui si decide di scaricare e utilizzare un'APP.
Da un sondaggio condotto dal *Pew Research Center* sugli utenti statunitensi è emersa una forte consapevolezza sulle informazioni richieste dalle APP, visto che il 90% degli intervistati ha sottolineato che tali informazioni sono molto importanti nella scelta di effettuare o no il *download* di un'applicazione, e che il 60% ha preferito non scaricare una APP dopo aver scoperto che i dati richiesti, in molti casi, non erano necessari al buon funzionamento dell'APP. Tuttavia, gli studi empirici sulla tutela della privacy hanno mostrato una serie di risultati contrastanti riguardo alla valutazione che gli utenti danno alla raccolta delle informazioni da parte delle APP. Questo aspetto sarà affrontato in maniera più approfondita nel capitolo successivo.

Nei capitoli precedenti, invece, si è ripetuto più volte che il valore dei dati varia a seconda delle caratteristiche del singolo dato e della combinazione con altri dataset di dati. Come i dati, anche i permessi, ovviamente, si differenziano tra loro.
Il *Pew Research Center*, in una ricerca condotta su un campione di APP disponibili su *Google Play Store*, ha distinto i permessi in due categorie: quelli riguardanti l'hardware e quelli riguardanti i dati degli utenti (cd. *user info*).
Una seconda importante classificazione è quella elaborata dai ricercatori Kummer e Schulte (2016). I due studiosi hanno individuato un numero limitato di permessi che possono risultare critici in termini di accesso a dati sensibili.
Facendo riferimento ai 10 permessi più diffusi tra le app, la **Figura 17** sintetizza, oltre al loro significato, anche se tali permessi possono essere considerati "sensibili" secondo le classificazioni del *Pew Research Center* e di Kummer & Schulte e se sono considerati

*dangerous* o *normal* sulla base della categorizzazione data direttamente dal gestore dello *store*, *Google.*

| Permesso | Il permesso consente all'APP di… | Classificazione | | |
|---|---|---|---|---|
| | | Pew Center | Kummer & Schulte | *Google* |
| Accesso completo a Internet (*full network access*) | …creare prese di rete e utilizzare protocolli di rete personalizzati. | Si | No | Dangerous |
| Vedere connessioni di rete (*view network connections*) | …vedere informazioni sulle connessioni di rete, ad esempio quali reti esistono e sono connesse | No | No | Normal |
| Accesso a memorie protette (*read the contents of your USB storage*) | …testare l'accesso a memorie USB o SD card che saranno disponibili su futuri device. | Si | No | Normal |
| Modificare e cancellare i contenuti della propria memoria USB (*modify or delete the contents of your USB storage*) | …sovrascrivere la memoria USB o la SD card. | Si | No | Dangerous |
| Leggere lo status e l'identità del telefono (*read phone status and identity*) | …accedere a caratteristiche telefoniche del device, quali il numero di telefono e l'ID del *device*. | Si | Si | Dangerous |
| Prevenire lo standby del device (*prevent device from sleeping*) | ….prevenire lo standby/pausa del *device*. | No | No | Normal |
| Vedere connessioni Wi-Fi (*view WI-FI connections*) | ….vedere informazioni sulle reti Wi-Fi (se sono attivate o quali sono i nomi dei device connessi). | Si | No | Dangerous |
| Posizione esatta (*precise location GPS and network-based*) | …accedere alla posizione esatta usando il Global Positioning System (GPS) o le risorse dela rete come le torri cellulari e il Wi-Fi. | Si | Si | Dangerous |
| Controllo della vibrazione (*control vibration*) | …controllare la vibrazione del *device*. | No | No | Normal |
| Posizione approssimativa (*approximate location network-based*) | …accedere alla posizione approssimativa usando le risorse della rete come le torri cellulari e il Wi-Fi. | Si | Si | Dangerous |
| **TOTALE PERMESSI "SENSIBILI"** | | **49** | **53** | **16** |

***"Figura 17: classificazione dei permessi"***

### 4.4.1 Correlazione tra concentrazione del mercato e permessi

Kesler et al., analizzando dati su oltre 1,5 milioni di app del Google Play Store, osservati su base trimestrale per più di due anni, da ottobre 2015 fino a gennaio 2018, sono riusciti a documentare una solida correlazione positiva tra concentrazione del mercato e raccolta dei dati, utilizzando l'analisi econometrica descritta di seguito.

$$Data_i = \alpha + \beta MP_i + \gamma Demand_i + \theta X_i + \varepsilon_i$$

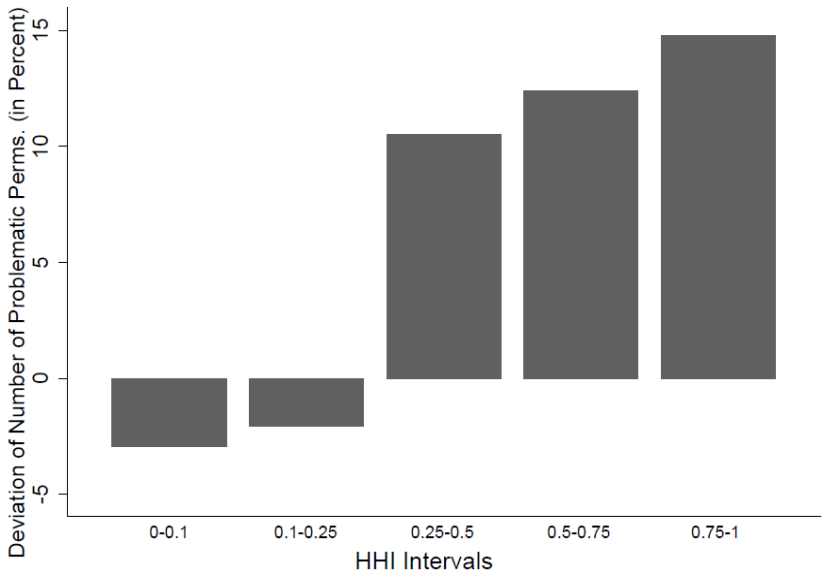
- $\boldsymbol{Data_i}$ è la quantità di dati raccolti per utente dall'app *i* ed è misurata dal numero di permessi problematici richiesti da un'app al momento dell'installazione.

- $\boldsymbol{Demand_i}$ è la domanda propria delle app, che è determinata dal numero (log) di valutazioni che un'app ha ricevuto.

- $\boldsymbol{X_i}$ sono una serie di variabili esplicative rilevanti per la raccolta dei dati.

- $\boldsymbol{MP_i}$ è il potere di mercato che viene misurato tramite:
    - concentrazione del mercato, utilizzando l'indice Herfindahl-Hirschman (HHI);
    - la quota di mercato logaritmica (Log. Market Share).

La **Figura 18** visualizza la relazione tra potere di mercato e raccolta dei dati, mostrando che le app attive in mercati altamente concentrati (HHI> 0,75) richiedono in media circa il 15% in più di autorizzazioni sensibili alla privacy rispetto all'app media complessiva, mentre app meno concentrate (HHI <0,1) richiedono il 2,5% in meno di permessi problematici rispetto all'app media complessiva.
La concentrazione del mercato è più elevata nelle categorie di Gioco e Comunicazione, mentre è più bassa nei gruppi di categorie per Meteo, Arte e Design. Ciò che ne consegue sarà quindi una maggiore raccolta dati in app come Facebook, Google gmail e altre app di gioco collegate da vicino al mondo Google.




*"Figura 18: relazione tra potere di mercato e raccolta dati"*

### 4.4.2 Correlazione download-permessi

Da analisi empiriche condotte su milioni di applicazioni da parte di AGCOM, emerge un rilevante effetto del sistema dei permessi sulle scelte dei consumatori (download).

In particolare, le relazioni che sono emerse evidenziano come il sistema dei permessi sia lo strumento attraverso il quale sono scambiati dati tra imprese e consumatori.

Alcune recenti ricerche hanno evidenziato la presenza di un nesso causale tra il numero di permessi richiesti da una APP e il comportamento degli utenti.

Gli studi empirici hanno posto attenzione sull'analisi della domanda di APP, procedendo alla stima di una regressione lineare che modella la domanda di APP, in termini di *download*, come funzione dei permessi che essa richiede, oltre che di una serie di variabili di controllo come il prezzo, la categoria di appartenenza, il *rating* medio e il numero di *review* che l'applicazione ha.

La specificazione del modello stimato è la seguente:

$$Domanda_i = \alpha + \beta D_i + \theta X_i + \varepsilon_i$$

$\boldsymbol{Domanda_i}$ = rappresentata dal logaritmo del numero totale di *download* della generica APP *i*;

$\mathbf{X_i}$ = include una serie di variabili di controllo (il prezzo, il totale dei permessi richiesti, il *rating* medio, il numero di *review* e categoria, lo sviluppatore dell'APP);

Il parametro $\boldsymbol{\beta}$ è quello di maggiore interesse poiché è associato a una variabile $\boldsymbol{D_i}$ *dummy* (ossia binaria) che tiene conto del fatto che un'APP richieda o no il consenso ad un permesso sensibile ai dati individuali. Di conseguenza, se il parametro stimato risulta negativo, ciò implica che la presenza di permessi sensibili alle informazioni individuali determina una riduzione nella domanda.

Per la stima del modello è stato utilizzo un semplice OLS i cui risultati sono esposti nella **Figura 19**.

Ciò che si osserva è una chiara relazione significativa tra il numero di download e i permessi riguardanti informazioni sensibili. In particolare, un aumento dei permessi, considerati *user info* secondo la classificazione del *Pew Research*, comporta una riduzione del 5% dei *download*.

Inserendo una classificazione più di dettaglio, come quella proposta da Kummer & Schulte, gli effetti dei permessi sono più contrastanti. Da un lato, infatti, quelli che richiedono l'accesso completo alla rete e al *device*, hanno un impatto positivo sui *download*; dall'altro lato, i permessi riguardanti le attività di comunicazione e la localizzazione dell'utente impattano, invece, in maniera negativa sul numero di *download*.

| | Modello con classificazione permessi *Pew Research* | Modello con classificazione permessi *Google* | Modello con permessi *Hummer e Schulte* |
|---|---|---|---|
| **Variabili di interesse** | | | |
| Permessi *user information* (PEW) | -0.05*** | | |
| | (0.00) | | |
| Permessi *dangeorus* (Google) | | -0.01*** | |
| | | (0.00) | |
| Permessi che consentono l'accesso alla rete | | | 0.07*** |
| | | | (0.00) |
| Permessi che consentono di vedere la presenza di reti | | | -0.13*** |
| | | | (0.00) |
| Permessi che consentono l'accesso allo stato del *device* | | | 0.07*** |
| | | | (0.00) |
| Permessi sulla localizzazione dell'utente | | | -0.05*** |
| | | | (0.00) |
| Permessi attinenti all'attività di comunicazione | | | -0.10*** |
| | | | (0.00) |
| Permessi riguardanti il profilo dell'utente | | | -0.02*** |
| | | | (0.00) |
| Altri permessi | | | -0.06*** |
| | | | (0.01) |
| Variabili di controllo | Sì | Sì | Sì |
| Categorie | Sì | Sì | Sì |
| F | 126634.03 | 126441.59 | 111810.44 |
| R-quadro | 0.84 | 0.84 | 0.84 |
| N | 1,135,700 | 1,135,700 | 1,135,700 |

Errori robusti all'eteroschedasticità in parentesi. ***, **, * coefficienti significativamente differenti da 0 all'1%, al 5% e al 10%

*"Figura 19: stime OLS modello di domanda di app"*

### 4.4.3 Correlazione prezzi-permessi

Un altro interessante risultato che è emerso da analisi empiriche, sempre effettuate da AGCOM, è che le APP gratuite richiedono un numero significativamente maggiore di permessi rispetto a quelle a pagamento. Vi è, quindi, un rilevante effetto del sistema dei permessi anche sui modelli di business che le imprese intendono adottare.

Per quel che riguarda l'analisi dell'offerta delle app, si è proceduto alla stima di un modello probabilistico in cui, tramite una variabile dicotomica uguale a 1 se un'APP è a

pagamento e uguale a 0 se un'app è gratuita, si analizza la scelta del modello di *business* da parte degli sviluppatori.

Il modello stimato è il seguente:

$$\Pr(\text{Prezzoi} = 1) = \Lambda[\alpha + \beta D_i + \theta X_i + \varepsilon_i]$$

Anche in questo caso, il principale parametro di interesse è rappresentato dal parametro $\boldsymbol{\beta}$ abbinato alla variabile *dummy* $\boldsymbol{D_i}$ che individua se un permesso è da considerarsi o meno rilevante rispetto ai dati individuali.

Al crescere del numero di permessi richiesti dalla APP, la probabilità che l'applicazione presenti un prezzo maggiore di 0 si riduce.

La **Figura 20** riporta i risultati delle stime del modello probabilistico, in cui vengono riportate le precedenti classificazioni dei permessi.

| | Modello con classificazione permessi *Pew Research* | Modello con classificazione permessi *Google* | Modello con permessi *Hummer e Schulte* |
|---|---|---|---|
| **Variabili di interesse** | | | |
| Permessi *user information* (PEW) | -0.26*** | | |
| | (0.00) | | |
| Permessi *dangeorus* (Google) | | -0.67*** | |
| | | (0.00) | |
| Permessi che consentono l'accesso alla rete | | | -0.41*** |
| | | | (0.01) |
| Permessi che consentono di vedere la presenza di reti | | | -0.55*** |
| | | | (0.00) |
| Permessi che consentono l'accesso allo stato del *device* | | | 0.09*** |
| | | | (0.00) |
| Permessi sulla localizzazione dell'utente | | | -0.30*** |
| | | | (0.01) |
| Permessi attinenti all'attività di comunicazione | | | 0.01 |
| | | | (0.01) |
| Permessi riguardanti il profilo dell'utente | | | -0.00 |
| | | | (0.01) |
| Altri permessi | | | -0.03 |
| | | | (0.02) |
| Variabili di controllo | Sì | Sì | Sì |
| Categorie | Sì | Sì | Sì |
| Pseudo R2 | 0.14 | 0.16 | 0.21 |
| N | 1,135,700 | 1,135,700 | 1,135,700 |

Errori robusti all'eteroscedasticità in parentesi. ***, **, * coefficienti significativamente differenti da 0 all'1%, al 5% e al 10%

*"Figura 20: stime probabilistiche del modello di offerta di app"*

Le APP a pagamento presentano, quindi, mediamente un numero di permessi inferiore rispetto a quelle gratuite (in media, circa la metà, ossia 3,8 a fronte di 6,4). In maniera implicita, dunque, il mercato attribuisce un valore economico ai dati digitali dell'utente: un loro minore rilascio, infatti, presuppone un prezzo più alto da pagare per l'utente.

| | n. di APP | % sul totale | n. medio di permessi |
|---|---|---|---|
| gratuite | 977.244 | 86% | 6,4 |
| a pagamento | 158.456 | 14% | 3,8 |
| **Totale** | **1.135.700** | **100%** | **6,0** |

Fonte: Elaborazioni AGCOM su dati *Google Play*

***"Figura 21: numero medio di permessi"***

# 5. Privacy

## 5.1 Definizione

Il concetto di privacy è un concetto *astratto* che non appartiene strettamente alla sfera economica. Il termine privacy può essere inteso come: la protezione dello spazio personale dell'individuo, il controllo e la salvaguardia delle informazioni personali, un aspetto di dignità, autonomia e libertà dell'individuo.

Un secolo fa, la risorsa più preziosa era il petrolio. Oggi nell'era digitale, i dati personali, o meglio la capacità di usarli al meglio e di combinarli tra di loro, costituiscono il nuovo petrolio, grazie ad una società che basa sul data driven l'intera sua esistenza.

Il governo di tale risorsa non risiede, però, nelle mani di un potere pubblico, bensì in quelle di giganti privati: Google, Amazon, Apple, Facebook, Microsoft sono le 5 aziende quotate più pregiate al mondo, con profitti inarrestabili, capaci di creare posizioni dominanti nei propri settori. Queste multinazionali digitali padroneggiano il nostro stato di salute, le nostre credenze religiose, le tendenze sessuali, insieme alle opinioni politiche e ai più intimi desideri di ciascuno.

Ogni volta che "clicchiamo" qualcosa sul web o effettuiamo una scelta su un dispositivo cediamo alla rete inconsapevolmente un pezzo di noi: un'opinione, un'emozione, un desiderio o un sogno. Come in un grande puzzle, la rete colleziona i singoli frammenti e, like dopo like, con meticolosa pazienza algoritmica, riesce a comporre una fotografia del nostro essere.

Nella pubblicità online, le reti pubblicitarie raccolgono grandi quantità di informazioni sui singoli utenti, che poi utilizzano per selezionare gli annunci da far visualizzare agli utenti mentre navigano sul web. Questo utilizzo dei dati rende gli annunci più pertinenti e informativi per l'utente, ma solleva anche problemi di privacy. Gli utenti non hanno un modo semplice per impedire la raccolta dei propri dati e non hanno alcuna garanzia che non saranno condivisi con altri attori che potrebbero usarli in modo dannoso contro di loro. Questa situazione è una rottura rispetto al passato, quando le discussioni pubbliche e legali si concentravano sulla raccolta e sull'utilizzo dei dati da parte del governo per la sorveglianza, la prevenzione della criminalità e l'individuazione del crimine.

Per le entità non governative, il discorso legale si è storicamente focalizzato sui casi in cui le imprese violavano la privacy, pubblicizzando informazioni personali e private sui personaggi pubblici. In passato, la raccolta di dati personali dettagliati era così costosa e

difficile che solo le persone che godevano di qualche forma di celebrità erano vulnerabili all'intrusione della privacy.
I recenti progressi tecnologici hanno reso la raccolta dei dati così scalabile che i dati di chiunque possono adesso essere raccolti e usati per scopi commerciali. I costi di raccolta e archiviazione dei dati sono scesi a un punto in cui quasi tutti hanno un interesse commerciale sufficiente a giustificare un monitoraggio elettronico.
È questo, dunque, il prezzo che dobbiamo corrispondere per vivere da cittadini digitali: spogliarci dei nostri dati, mettendo a nudo la nostra vera personalità nelle mani delle piattaforme di internet, che di contro non ci restituiscono nessuna trasparenza su come tali dati vengono impiegati.

"Cittadino digitale" nel proprio senso della parola è ciò a cui sta pensando Google che sta progettando un'intera città gestita in modo intelligente. Si chiamerà Quayside e si preannuncia già rivoluzionaria, la prima città intelligente e completamente automatizzata, gestita dagli algoritmi che regoleranno il traffico, consentiranno di risparmiare energia, ridurranno l'inquinamento e avvertiranno anche quando è ora di svuotare i bidoni della spazzatura. Il primo quartiere governato interamente da Internet nascerà a Toronto, in una vecchia area industriale dismessa. Grazie alla comunicazione di milioni di sensori disseminati in ogni angolo di Google city, tutti gli abitanti saranno costantemente osservati per offrire loro tutto il potenziale dell'innovazione tecnologica, un vero e proprio "abito su misura" grazie a chip, antennine e big data. I semafori si regoleranno automaticamente, percependo la presenza di auto e pedoni, gli edifici si oscureranno in presenza di forti giornate di sole, così da donare refrigerio ai propri inquilini, le strade saranno riscaldate e capaci di sciogliere la neve, le auto saranno esclusivamente a guida autonoma.
I punti di criticità sono evidenti. Questo nuovo ecosistema digitale imporrà un notevole uso di dati personali: gli spostamenti dei cittadini dovranno essere costantemente registrati, preferenze, usi e abitudini degli abitanti saranno incessantemente scambiati dai vari dispositivi. L'imponente mole di informazioni e dati sensibili rappresenterà inoltre una merce ambitissima per tutti i servizi offerti dal motore di ricerca, che potrebbe facilmente usarli per ulteriori fini, e non soltanto per gestire il centro abitato.

La privacy è quindi un problema rilevante e urgente dei nostri tempi. Se il progresso tecnologico sembra sempre più affascinare e correre a dei ritmi futuristici, la tutela dei diritti sulla privacy, invece, non riesce a stare al passo.
Oggi, la parola ha lasciato spazio alla navigazione. Parliamo di meno e scriviamo di più, trascurando il valore di una chiacchierata telefonica e producendo tonnellate di dati sotto forma di messaggi, audio e pagine web. Riveliamo informazioni sia consapevolmente che inconsapevolmente e il monitoraggio è onnipresente; la sua conservazione è così duratura da rendere il passato non cancellabile, un moderno scheletro digitale nell'armadio.
Queste nuove dinamiche comportamentali non fanno altro che aumentare la produzione di migliaia di informazioni, tracce sulle proprie abitudini, preferenze, credenze, opinioni, mettendo così in bella vista la nostra privacy.

Alcuni studiosi si fidano della razionalità, dell'intelligenza e quindi della capacità delle persone di prendere decisioni, riguardo alla divulgazione e alla ritenuta delle informazioni.
Coloro che sostengono questo punto di vista tendono a vedere la protezione normativa della privacy come un ostacolo alla traiettoria fondamentalmente benigna delle tecnologie dell'informazione e ai vantaggi che tali tecnologie potrebbero sbloccare.
Altri, invece, sono preoccupati per la capacità delle persone di gestire la privacy in mezzo a compromessi sempre più complessi. Gli strumenti tradizionali per il processo decisionale in materia di privacy come la scelta e il consenso, secondo questa prospettiva, non forniscono più una protezione adeguata. La raccolta dei dati da parte delle piattaforme può portare a degli abusi, ovvero a dei comportamenti che possono danneggiare l'utente, ora o nel futuro. Queste problematiche sorgono proprio dalla caratteristica di replicabilità dei dati: è molto semplice replicare un dato e venderlo a terzi, ciò aumenta il trasferimento e la circolazione dei dati.
In particolare, possiamo identificare tre tipologie di problematiche legate alla privacy:

- ***Data Persistence***: il dato, una volta creato, può sopravvivere molto più a lungo rispetto al motivo per il quale l'utente l'ha fornito. Ciò è dovuto al fatto che il costo di storage dei dati è molto basso.

- ***Data Repurposing***: non è chiaro come il dato possa essere utilizzato in futuro, ma c'è il rischio che possa essere utilizzato per danneggiare lo stesso utente che lo ha fornito.

- ***Data Spillovers***: i dati da me forniti possono avere effetti anche per altri utenti che non ne hanno voluto fornire.

Il problema del controllo sui dati personali è diventato, quindi, legato a problemi di scelta personale, autonomia e potere socioeconomico.

## 5.2 Dipendenza dal contesto

Applicata alla privacy, la dipendenza dal contesto significa che le persone possono, a seconda della situazione, esibire qualsiasi cosa, dall'estrema preoccupazione all'apatia per la privacy.
Adottando la terminologia di Westin, siamo tutti pragmatici della privacy, fondamentalisti della privacy o indifferenti, a seconda del tempo e del luogo.
Le regole che le persone seguono per la gestione della privacy variano in base alla situazione, sono apprese nel tempo e si basano su criteri culturali, motivazionali e puramente situazionali.
Le norme e i comportamenti riguardanti i regni privato e pubblico differiscono notevolmente tra le culture. Gli americani, ad esempio, sono ritenuti più aperti sulle questioni sessuali rispetto ai cinesi, mentre questi ultimi sono più aperti sulle questioni finanziarie (come il reddito, il costo della casa e le proprietà).
Le preoccupazioni sulla privacy possono variare notevolmente per la stessa persona e per le società nel tempo. Ad esempio, di solito potremmo essere più a nostro agio nel condividere segreti con gli amici, ma a volte possiamo rivelare informazioni sorprendentemente personali anche a uno sconosciuto su un aereo.
È stato, inoltre, dimostrato che la presenza di normative governative riduce le preoccupazioni dei consumatori e aumenta la fiducia.

Un altro aspetto che influenza la percezione della privacy è il comportamento di altre persone. Osservare le altre persone rivelare informazioni aumenta la probabilità che le si rivelino da sole. In uno studio, ai partecipanti al sondaggio sono state poste una serie di domande personali delicate riguardanti il loro impegno in comportamenti illegali o

eticamente discutibili. Dopo aver risposto a ciascuna domanda, ai partecipanti sono state fornite informazioni, manipolate a loro insaputa, sulla percentuale di altri partecipanti che nella stessa indagine avevano ammesso di aver assunto un determinato comportamento. Essere forniti di informazioni che suggerivano che la maggior parte dei partecipanti al sondaggio aveva ammesso un certo comportamento discutibile aumentava la volontà dei partecipanti di svelare altri loro comportamenti, anche sensibili.

Altri studi hanno scoperto che la tendenza a ricambiare la divulgazione delle informazioni è così radicata che le persone rivelano più informazioni anche a un agente informatico che fornisce informazioni su se stesso. Risultati come questo possono aiutare a spiegare le crescenti quantità di auto-divulgazione a cui assistiamo online: se gli altri lo fanno, le persone sembrano ragionare inconsciamente, pensando quindi che condividere le proprie informazioni debba essere desiderabile e sicuro anche per loro.

Gli esseri umani sono animali sociali e la condivisione delle informazioni con altre persone è una caratteristica centrale della connessione umana. Il desiderio di interazione, socializzazione, divulgazione sono motivi umani non meno fondamentali del bisogno di privacy. Attraverso i social media, la divulgazione può costruire capitale sociale, aumentare l'autostima e soddisfare i bisogni dell'ego. In una serie di esperimenti di imaging a risonanza magnetica funzionale, si è scoperto che l'auto-divulgazione coinvolge meccanismi neurali associati alla ricompensa; le persone apprezzano molto la capacità di condividere pensieri e sentimenti con gli altri.

I problemi di privacy sono spesso anche una funzione delle esperienze passate. Quando qualcosa in un ambiente cambia, come ad esempio l'introduzione di una videocamera o di altri dispositivi di monitoraggio, è probabile che si attivino problemi di privacy. Ad esempio, la sorveglianza può provocare disagio e influire negativamente sulla produttività dei lavoratori. Tuttavia, le preoccupazioni sulla privacy, come altre motivazioni, sono adattive. Le persone, infatti, si abituano a nuovi livelli di intrusione. In un esperimento condotto a Helsinki, l'installazione della tecnologia di rilevamento e monitoraggio nelle famiglie ha portato inizialmente i membri della famiglia a cambiare il loro comportamento, in particolare in relazione a conversazioni, nudità e sesso. Eppure, se eseguissero accidentalmente un'attività, come camminare nudi in cucina davanti ai sensori, sembrerebbe avere l'effetto di "rompere il ghiaccio" e i partecipanti mostrerebbero meno preoccupazione per la ripetizione del comportamento.

## 5.3 Malleabilità e influenza

Con l'emergere dell'era dell'informazione, si sono sviluppati crescenti interessi istituzionali ed economici sulla divulgazione di informazioni personali, dai social network online alla pubblicità comportamentale.

Non sorprende, quindi, l'interesse da parte di molte imprese nello sviluppo di competenze per sfruttare al meglio i processi comportamentali e psicologici e per promuovere la divulgazione. Tali sforzi giocano sulla malleabilità delle preferenze sulla privacy.

Le impostazioni predefinite sono uno strumento importante, adoperato da diversi attori per influenzare la divulgazione delle informazioni personali. Attenersi alle impostazioni predefinite è conveniente e le persone spesso interpretano le impostazioni predefinite come raccomandazioni implicite, non dandole la benché minima attenzione. Pertanto, non sorprende che le impostazioni predefinite per la visibilità del proprio profilo sui social network influenzino il comportamento della privacy delle persone.

Oltre alle impostazioni predefinite, i siti Web possono anche utilizzare funzionalità di progettazione che confondono gli utenti nella divulgazione di dati sensibili, una pratica che è stata definita "progettazione di interfacce dannose".

Il controllo sulle informazioni personali è spesso percepito come una caratteristica fondamentale della protezione della privacy. In linea di principio, fornisce agli utenti i mezzi per gestire l'accesso alle loro informazioni personali. La ricerca, tuttavia, mostra che il controllo può ridurre la preoccupazione per la privacy, che a sua volta può avere effetti indesiderati. Ad esempio, uno studio ha scoperto che i partecipanti a cui veniva fornito un maggiore controllo esplicito sulla raccolta dei dati finivano per condividere informazioni più sensibili con un pubblico più ampio, l'opposto dell'obiettivo apparente di fornire tale controllo.

La ricerca ha evidenziato, inoltre, non solo che la stragrande maggioranza degli utenti di Internet non legge le politiche sulla privacy, ma anche che pochi utenti ne trarrebbero beneficio. Quasi la metà di un campione di politiche sulla privacy online è stato trovato in un linguaggio che va oltre la portata della maggior parte degli utenti di Internet.

## 5.4 Paradosso della privacy

Gli individui gestiscono i confini tra le loro sfere pubbliche e private in numerosi modi: tramite separazione, riservatezza o anonimato, proteggendo le informazioni personali, ma anche attraverso l'inganno. Le persone stabiliscono tali limiti per molte ragioni, tra cui la necessità di intimità e il desiderio di protezione dall'influenza e dal controllo sociale.
I progressi nella tecnologia dell'informazione hanno reso spesso invisibili la raccolta e l'utilizzo di dati personali. Di conseguenza, le persone raramente hanno una chiara conoscenza di quali informazioni altre persone, aziende e governi hanno su di loro o come tali dati sono utilizzati e con quali conseguenze. Anche quando sono consapevoli delle conseguenze delle decisioni sulla privacy, le persone sono incerti sulle proprie preferenze sulla privacy.

Westin ha usato ampie domande nei sondaggi per raggruppare gli individui in segmenti di privacy: fondamentalisti della privacy, pragmatici e indifferenti. Quando sono stati interrogati direttamente, poche persone che rientravano nel primo segmento (professavano di preoccuparsi molto della privacy e di esprimere particolare preoccupazione per la perdita di controllo delle loro informazioni personali) hanno mostrato riluttanza a rispondere alle domande sempre più delicate a cui l'agente li ha sottoposti. Il 48% di coloro che avevano espresso il massimo grado di preoccupazione ha, infatti, rivelato pubblicamente il loro orientamento sessuale online, il 47% ha rivelato il loro orientamento politico e il 21% ha rivelato il nome del loro attuale partner.
Questa discrepanza tra atteggiamenti e comportamenti è diventata nota come il **"paradosso della privacy"**. Spesso, i consumatori dichiarano una forte preferenza per la privacy, ma sono disposti a rinunciare alle proprie informazioni personali per piccoli incentivi. Si potrebbe avere a cuore la privacy in generale ma, a seconda dei costi e dei benefici prevalenti in una situazione specifica, cercare o meno la protezione.
Il processo decisionale sulla privacy, tuttavia, è solo in parte il risultato di un "calcolo" razionale di costi e benefici. È inoltre influenzato, infatti, da errate percezioni di tali costi e benefici, nonché da norme sociali, emozioni ed euristica. Ognuno di questi fattori può influenzare il comportamento in modo diverso da come influenzano gli atteggiamenti.

## 5.5 Monetising privacy

L'incertezza delle preferenze è evidente non solo negli studi che confrontano gli atteggiamenti dichiarati con i comportamenti, ma anche in quelli che stimano le valutazioni monetarie della privacy.

L'economia della privacy è un campo di ricerca all'incrocio tra economia, diritto e informatica. È dedicato allo studio dei compromessi economici costi-benefici che le persone intraprendono quando divulgano i dati personali nelle transazioni economiche (il cosiddetto “calcolo della privacy”).

### 5.5.1 Valore del dato come disponibilità ad accettare degli incentivi monetari

Indagini "esplicite" hanno chiesto alle persone di effettuare scambi diretti, in genere tra privacy dei dati e denaro. Ad esempio, in uno studio condotto, da Hann e altri ricercatori, sia a Singapore che negli Stati Uniti, gli studenti hanno fatto una serie di ipotetiche scelte sulla condivisione di informazioni, con siti Web che differivano nella protezione delle informazioni personali e dei prezzi per l'accesso ai servizi. Le informazioni personali consistevano in nome, genere, età, indirizzo di casa, numero di telefono, indirizzo e-mail, informazioni sulla carta di credito e alcune informazioni specifiche del settore. In particolare, i siti Web di viaggio richiedevano l'occupazione della persona, lo scopo del viaggio, la destinazione e la frequenza del viaggio; i siti Web sanitari chiedevano anamnesi, allergie ai farmaci e record di prescrizione; e siti Web finanziari chiedevano reddito familiare, portafoglio azionario e precedenti esperienze di trading azionario. I soggetti hanno valutato la protezione da errori, accesso improprio e uso secondario delle informazioni personali tra $ 30,49 e $ 44,62.

I siti Web, quindi, dovrebbero offrire incentivi monetari importanti per superare le preoccupazioni degli individui riguardo alla propria privacy.

Huberman et al. (2005) hanno progettato un esperimento che ha permesso di estrarre il valore che ogni individuo attribuisce alla divulgazione delle informazioni private, in particolare peso ed età. Il risultato principale è che meno le informazioni sul peso o sull'età rivelate erano socialmente desiderabili (rispetto alla media del gruppo), maggiore era il prezzo richiesto da una persona per pubblicarle.

Gli individui di peso leggermente inferiore alla media non hanno richiesto molti soldi per pubblicizzarlo, in media solo $ 2,71. D'altra parte, quelli che pesavano di più e che potevano temere imbarazzo o stigmatizzazione chiedevano molto di più, in media $ 41,32
Per quanto invece riguarda l'età, i soggetti molto giovani sono più disposti a rivelare la loro età rispetto ai più anziani. Tra i due contenitori estremi troviamo una differenza di prezzo significativa ($ 3,62 contro $ 18,05).
Nel sondaggio è stato anche chiesto ai partecipanti di indicare quanto avrebbero richiesto per altre tipologie di dati come lo stipendio, lo stipendio coniugale, il rating del credito e i risparmi. La percentuale di individui che chiedevano più di $ 100 era rispettivamente del 48% per lo stipendio, 36% per lo stipendio del coniuge, 24% per la solvibilità e 38 % per il risparmio. Tutti sono relativamente alti rispetto a peso ed età in cui solo il 5,5% e il 3,5% rispettivamente ha chiesto più di $ 100. Questo costo aggiuntivo può essere in parte correlato a tabù sociali che impediscono la discussione aperta di informazioni come lo stipendio, al fine di prevenire conflitti.

### 5.5.2 Valore del dato come disponibilità a pagare

Sebbene le persone a volte rinuncino ai dati personali per benefici o sconti, altre volte comportano volontariamente costi sostanziali per proteggere la loro privacy.
In Tsai et al. (2010), gli autori hanno scoperto che quando le informazioni sulla politica sulla privacy sono visualizzate in modo più saliente, i partecipanti prendono maggiormente in considerazione la politica sulla privacy e tendono ad acquistare da rivenditori online che ottengono un punteggio più alto nell'indice di protezione e tutela dei dati. In questo caso, sono persino inclini a pagare un premio per i siti Web che proteggono meglio la loro privacy.

Beresford et al. (2010) hanno utilizzato un esperimento per analizzare la disponibilità a pagare per la privacy, in cui ai partecipanti veniva offerta la possibilità di acquistare un DVD da uno dei due negozi online concorrenti. Sebbene questi negozi fossero identici, uno richiedeva dati personali più sensibili dell'altro.
Nel trattamento di prova, quando i DVD costavano € 1 in meno presso l'azienda invasiva della privacy, praticamente tutti gli acquirenti hanno scelto il negozio più economico.
Nel trattamento di controllo con prezzi identici, le persone non preferivano sistematicamente l'azienda più rispettosa della privacy, ma sceglievano entrambe le

aziende allo stesso modo. L'esperimento ha dimostrato una riluttanza a pagare per la privacy poiché la stragrande maggioranza dei soggetti fornisce il proprio reddito mensile con uno sconto di un solo euro. Anche senza uno sconto sul prezzo, solo la metà dei soggetti ha effettuato acquisti presso il ramo più rispettoso della privacy del rivenditore di DVD. Questo risultato è sorprendente dato che la maggior parte dei soggetti che forniscono informazioni sensibili non ne è soddisfatta.

In ENISA, "Study on monetising privacy. An economic model for pricing personal information" (2012), gli autori hanno effettuato il più grande esperimento di laboratorio nel campo dell'economia della privacy fino ad oggi. Sono stati implementati diversi scenari, in cui 443 partecipanti si sono trovati di fronte a due diversi fornitori di servizi che offrivano biglietti per il cinema. Entrambe le aziende nell'esperimento hanno sempre chiesto un set minimo di dati personali come nome, cognome, indirizzo e-mail e data di nascita.
L'esperimento di laboratorio ha mostrato le seguenti conclusioni:

- una percentuale non trascurabile dei partecipanti all'esperimento (13,83%) ha scelto di pagare un "premio" per la privacy;

- il fornitore di servizi rispettoso della privacy ha una quota di mercato molto più ampia, se le differenze nella raccolta dei dati sono evidenti e i prezzi sono gli stessi;

- se un concorrente ostile alla privacy addebita un prezzo inferiore, il fornitore di servizi rispettoso della privacy perde quote di mercato.

Se confrontiamo lo scenario in cui le imprese sono identiche (Scenario 1) con la situazione in cui variano in base al numero di articoli che raccolgono (Scenario 2), variando solo quest'ultimo stimolo, troviamo che la quota di mercato dell'Impresa 1, l'impresa rispettosa della privacy, è significativamente più alta nel secondo scenario rispetto al primo e che tale differenza è statisticamente significativa.
Se non ci sono differenze di prezzo e differenze nei requisiti di dati (Scenario 1), viene raccolto oltre il 60% della quota di mercato in termini di acquisti per l'impresa 1. Aumenta, invece, all'83% se ci sono differenze nei requisiti dei dati (Scenario 2).

Se confrontiamo, inoltre, lo scenario in cui le imprese variano in base al numero di elementi di dati raccolti (scenario 2) con la situazione in cui variano in base a elementi sia di dati che prezzi (scenario 3), i risultati cambiano notevolmente. In quest'ultimo caso, l'Azienda 1, rispettosa della privacy, addebita 0,50 € in più rispetto al suo concorrente ostile alla privacy. La quota di biglietti venduti dall'azienda attenta alla privacy ora diminuisce fortemente (dall'83% al 31%) dallo scenario 2 a quello 3.
La differenza tra i trattamenti è statisticamente significativa. Ciò significa che la quota di mercato dell'azienda rispettosa della privacy è notevolmente ridotta, una volta che un concorrente addebita un prezzo inferiore di appena 0,50 €, raccogliendo al contempo più informazioni.

In Acquisti et al. (2013), è stata sottoposta un'altra indagine a degli acquirenti di un centro commerciale a cui sono stati offerti buoni regalo per la partecipazione a un sondaggio. Ai partecipanti è stata data una carta regalo "anonima" da $ 10 (le transazioni effettuate con quella carta non sarebbero state riconducibili all'oggetto) o una carta tracciabile da $ 12 (le transazioni fatte con quella carta sarebbero collegate al nome dell'oggetto).
Inizialmente, ad una metà dei partecipanti veniva assegnato un tipo di carta e l'altra metà all'altro tipo e in seguito, a tutti è stata offerta l'opportunità di cambiare.
Ad alcuni acquirenti, ad esempio, è stata data la carta anonima da $ 10 e gli è stato chiesto se avrebbero accettato $ 2 per consentire al nome di essere collegato alle transazioni effettuate con la carta; ad altri soggetti è stato chiesto se avrebbero accettato una carta con un valore inferiore di $ 2 per impedire che il nome fosse collegato a transazioni effettuate con la carta.
La manipolazione della dotazione della carta regalo ha generato una diversa inquadratura della scelta della carta: per quelli che possedevano la carta di $ 10, la domanda è stata inquadrata come una scelta implicita di vendere i propri dati ai ricercatori per degli incentivi monetari di $ 2; per quelli che possedevano la carta di $ 12, la domanda è stata inquadrata come una scelta implicita di pagare $ 2 per evitare che i propri dati fossero resi disponibili ai ricercatori.
Il 52,1% dei soggetti, che in origine possedevano la carta da $ 10, ha mantenuto la carta anonima, mentre il restante 47,9% ha accettato 2 $ in più e scambiato la carta.

Solo il 9,7% dei soggetti che avevano in origine la carta da $ 12 ha cambiato, scegliendo la carta anonima e pagando 2 $ per la propria privacy, mentre il restante 90,3% ha preferito non cambiare.

Ciò evidenzia questa dicotomia tra diponibilità ad accettare degli incentivi monetari come merce di scambio per i nostri dati e disponibilità a pagare per la nostra privacy, e suggerisce inoltre che, nella maggior parte dei casi, le persone apprezzano maggiormente la privacy quando ce l'hanno rispetto a quando invece no. Il valore che associamo ai nostri dati, infatti, è molto più alto quando li difendiamo da chi li vuole comprare, piuttosto al caso in cui dobbiamo pagare dei premi per difenderli.

### 5.5.3 Valore del dato oggettivo

Nel 2013, il Financial Times ha cercato di stabilire il valore dei dati personali attraverso l'utilizzo di un calcolatore online, creato sulla base dell'analisi dei dati sui prezzi del settore da una serie di fonti negli Stati Uniti.

Questo calcolatore, ovviamente, non è completo e non include i dettagli sui prezzi su ciascuna delle migliaia di bit di informazioni che i broker di dati tracciano, analizzano e vendono sugli individui, ma è comunque un algoritmo che cerca di approssimare in modo abbastanza preciso ciò che avviene in realtà.

L'algoritmo, inizialmente, ipotizza che età, genere, livello d'istruzione ed etnia siano disponibili. I valori approssimativi sono presentati in **Figura 22**.

| | |
|---|---|
| ETÀ | $ 0,0005 |
| GENERE | $ 0,0005 |
| LIVELLO DI ISTRUZIONE | $ 0,0005 |
| ETNIA | $ 0,005 |

*"Figura 22: valori dei principali dati raccolti"*

Successivamente, l'algoritmo pone una trentina di domande su quattro pagine e l'utente, rispondendo a questa serie di domande inerenti al proprio status (dai beni posseduti, agli hobby, passando per malattie croniche e stato familiare), può conoscere il valore commerciale di ogni singolo dato rilasciato ed incrementare il valore di quanto già ceduto. Una persona che dichiara di stare per acquistare un'auto, un prodotto finanziario o una vacanza è leggermente più preziosa per le aziende desiderose di lanciare quei beni.
Alcune pietre miliari nella vita di una persona provocano grandi cambiamenti nei modelli di acquisto, che si tratti di diventare un nuovo genitore, spostare case, fidanzarsi, acquistare un'auto e tutti questi dati sono molto importanti per le aziende che vogliono conoscerci sempre di più, in modo da offrirci il bene o il servizio che fa per noi.
Dichiarare per esempio di avere degli hobby come la corsa ha un valore di circa $ 0,08, possedere una casa ed essersi trasferiti nell'ultimo mese ha un valore di $ 0,2, possedere una barca circa $ 0,1.
Mediamente, secondo l'algoritmo, i dati di ogni persona valgono circa $ 0,6.

### 5.5.4 Confronto tra i tre valori

Un conteggio del valore effettivo dei nostri dati è quasi impossibile da stabilire, poiché troppi sono gli attori che beneficiano in un modo o nell'altro dei nostri dati.
Il valore del dato, come abbiamo visto dalle ricerche effettuate, dipende da innumerevoli fattori: influenza e malleabilità, dipendenza dal contesto, zona geografica, fattori culturali, desiderabilità sociale nel divulgare, asimmetrie informative.
Il valore, inoltre, oscilla molto a seconda del soggetto interessato: un VIP o una persona molto ricca che dichiara di possedere molti beni è più importante per le aziende e ha quindi più valore in quanto dotato di capacità di acquisto superiore, mentre i dati di una persona di ceto più basso valgono meno per il minor potere di spesa.
Nonostante queste difficoltà, è stato possibile osservare il valore del dato sotto tre differenti punti di vista:

- disponibilità ad accettare degli incentivi monetari;
- disponibilità a pagare dei premi;
- oggettivo (algoritmo del Financial Times).

| Valori associati a informazioni personali: nome, genere, età, indirizzo di casa, numero di telefono, indirizzo e-mail. | DISPONIBILITÀ AD ACCETTARE DEGLI INCENTIVI MONETARI | DISPONIBILITÀ A PAGARE | OGGETTIVO |
|---|---|---|---|
| **Hann, Il-Horn, Kai-Lung, "Overcoming Information Privacy Concerns: An Information Processing Theory", 2007** | [$ 30,49; $ 44,62] | | |
| **Huberman, B.A., Adar, E. and Fine, L.R. "Valuating Privacy", 2005** | [$ 2,71; $ 41,32] **(Solo Peso)** [$ 3,62; $ 18,05] **(Solo Età)** | | |
| **Beresford, A.R., Kübler, "Unwillingness to Pay for Privacy: A Field Experiment", 2010** | | € 1 | |
| **ENISA, "Study on monetising privacy. An economic model for pricing personal information", 2012** | | € 0,50 | |
| **Alessandro Acquisti, Leslie K. John, and George Loewenstein, "What Is Privacy Worth?", 2013** | >> € 2,00 | € 2,00 | |
| **Financial Times, 2013** | | | $ 0,6 $ 0,0005 **(Solo Età)** |

*"Figura 23: studi a confronto"*

Ciò che emerge da tutti questi studi in materia è che le nostre informazioni assumono un valore maggiore quando ci viene chiesto di venderle. Il valore che assegniamo ai nostri dati è molto alto e siamo disposti a venderli solo accettando degli incentivi monetari abbastanza elevati, come possiamo notare nella **Figura 23.**

Acquisti et al. (2013) evidenziano questa dicotomia tra diponibilità ad accettare degli incentivi monetari come merce di scambio per i nostri dati e disponibilità a pagare per la nostra privacy. Notano come una cifra di $ 2 possa già risultare elevata se siamo noi, utenti, a dover pagare per difendere la privacy, mentre insufficiente nel caso dovessimo accettare questa somma per poter vendere le nostre informazioni.

Questi due punti di vista anche se presentano dei valori molto differenti hanno in comune la soggettività nelle valutazioni monetarie della privacy.

Se ragionassimo in termini puramente economici, oggettivi, il valore dei dati si abbassa ulteriormente. Il valore oggettivo, assunto dall'algoritmo del Financial Times, infatti, è una cifra bassissima, perché il valore dei dati è, implicitamente per noi, qualcosa di molto soggettivo, ma se ragioniamo in termini economici, considerando la relativa facilità e imprudenza con cui svendiamo queste informazioni nei vari social, è un prezzo che cominciamo a valutare diversamente. Il prezzo lo fissa il mercato. La domanda di dati è importante; tuttavia, se l'offerta dei nostri dati è così eccessiva non c'è motivo che il prezzo sia alto.

Come ripetuto nei precedenti capitoli, raccogliere alcune tipologie di dati, oggigiorno, è relativamente semplice. Come veri e propri intermediari, i Data Broker o anche le stesse app sono entità che raccolgono informazioni online sui consumatori da fonti pubbliche, le aggregano, le interpretano e le analizzano per poi vendere quei dati ad altri data broker o aziende. Ciò, però, non ci deve far dedurre erroneamente che i dati abbiano poco valore. È proprio la combinazione dei dati ad attribuire maggior valore complessivo e le piattaforme come Facebook, Amazon, Netflix sono ovviamente avvantaggiate nel mercato della raccolta dati.

Il valore, quindi attribuito dall'algoritmo è solo un valore razionale, che si basa sulla facilità di come questi dati vengano ormai oggi rilasciati sul web, ma non considera a pieno il loro potenziale.

Se osserviamo, inoltre, più da vicino lo studio di Huberman e l'algoritmo del Financial Times possiamo notare come una medesima informazione generica come l'età assuma dei valori differenti nei due casi.

Siamo disposti a vendere un'informazione come la nostra età per un incentivo di $ 18, perché sappiamo che per noi è un'informazione importante e la vorremmo difendere, ma spesso la rilasciamo, senza neanche accorgercene, quando scarichiamo un'app di giochi sul nostro cellulare. È proprio per questo motivo che l'algoritmo del Financial times associa a questa informazione un valore di appena $ 0,0005.

Questo è un esempio lampante di ciò che vuol dire paradosso della privacy, in termini economici.

L’utente, quindi, prima mostra un atteggiamento forte a favore della propria privacy, dichiarandosi inamovibile, ma, in seguito, mostra un comportamento differente perché è disposto a cedere un’informazione sensibile per un incentivo monetario.
Infine, spesso, senza neanche accorgersene la rilascia gratuitamente o comunque non è molto disponibile a pagare dei premi per difenderla.

# 6. GDPR

La nuova economia dei dati ha evidenti benefici sia per le aziende che per le persone, ma solleva preoccupazioni sulla privacy. Mai prima d'ora le imprese sono state in grado di osservare le azioni dei consumatori a un livello così dettagliato o ottenere tali informazioni potenzialmente personali. Tali capacità generano la possibilità di una tensione tra innovazioni che si basano sull'uso dei dati e la protezione della privacy dei consumatori.

La questione della regolazione di Internet comincia a sollevare interrogativi di ordine etico a partire dalla sua stessa impostazione. Infatti, l'idea stessa del web è direttamente associata a spazio di libertà, la sua regolazione comporterebbe il venir meno del suo tratto distintivo, snaturandone l'essenza. Nello stesso tempo, tale idea di libertà assoluta non corrisponde più alla realtà.

Nata come strumento di libertà, ben presto la rete si è piegata agli interessi economici delle tech companies, organizzazioni che se da una parte hanno avuto il merito di contribuire all'espansione e all'aumento delle funzionalità del web, dall'altra hanno realizzato forme di controllo tanto sottili quanto invasive, realizzando un modello di sviluppo che, se non opportunamente regolato, può condurre alla manipolazione delle coscienze.

Occorre quindi attuare azioni per "governare la rete" che consistono, essenzialmente, in procedure, norme e regole aperte a tutti, e finalizzate a gestire sia gli aspetti prettamente tecnici, sia quelli etico-sociali, così da assicurare regole uniformi, poste a salvaguardia di principi irrinunciabili, anche nella sfera virtuale, come la neutralità della rete, la libertà di espressione, la tutela della privacy.

Il tema della protezione della privacy e dei dati personali è apparso spesso nelle riflessioni e nei dibattiti degli ultimi anni, soprattutto nel contesto del social networking e della profilazione degli individui. La presenza di poche qualificate piattaforme, capaci di veicolare quello che gli utenti pensano, scrivono e condividono, rende ancora più palese la contraddizione di Internet: da una parte assistiamo ad un processo straordinario di apertura nella produzione di informazioni, ma, dall'altro lato, la disponibilità di tali informazioni si concentrano nelle mani di pochi attori che decidono cosa mostrare, in quale posizione e per quanto tempo, minando alla base la libertà e il pluralismo dell'informazione stessa.

Le piattaforme, dopo aver raccolto i nostri dati, strettamente personali, sono anche in grado di filtrare le notizie, le informazioni disponibili e presentarle agli utenti secondo un ordine prestabilito, spesso personalizzato. Internet, quindi, non ottiene soltanto i nostri dati, le nostre preferenze, ma è capace anche di condizionarli.

C'è bisogno, quindi, di regole, prima fra tutte quelle relative alla riservatezza dei dati degli utenti.

In questo contesto, nell'aprile del 2016, l'Unione Europea è stata la prima a adottare il Regolamento generale sulla protezione dei dati (GDPR), che consente ai consumatori di avere un maggiore controllo sui dati che generano. Le società hanno avuto tempo due anni per conformarsi alle nuove normative e, il 25 maggio 2018, il GDPR è diventato realtà in tutte le aziende. Il GDPR stabilisce norme comuni sul trattamento dei dati, per garantire una protezione efficace dei dati personali, mantenendo i benefici del trattamento dei dati, in tutta l'Unione europea.

## 6.1 Le principali novità introdotte dal GDPR

Il Regolamento introduce una disciplina in materia di protezione dati uniforme e valida in tutta Europa. Essendo obbligatorio in tutti i suoi elementi e direttamente applicabile in ciascuno degli Stati membri, è definito "self-executing" in quanto le sue disposizioni sono immediatamente applicabili, senza il bisogno di alcuna intermediazione.

Gli effetti di questa "differenza sistemica" tra il Regolamento UE 2016/679 (GDPR) e le direttive precedenti non riguardano soltanto gli adempimenti richiesti agli Stati membri, i quali sono chiamati a adeguare le proprie normative interne al mutato regime europeo e garantirne la sua piena applicabilità, ma si riflettono principalmente sull'intera articolazione interna della disciplina. Con il Regolamento, e il suo assetto legislativo, incentrato sulla diretta applicabilità delle regole europee, si ha come naturale conseguenza una maggiore completezza e chiarezza dei contenuti normativi.

Il GDPR ha richiesto cambiamenti sostanziali nel modo in cui le aziende archiviano ed elaborano i dati dei consumatori. Le imprese, infatti, sono adesso tenute ad essere più esplicite sulla loro politica di conservazione dei dati, giustificando il periodo di tempo in cui conservano le informazioni sui consumatori. Inoltre, il GDPR ha richiesto alle imprese di aumentare la trasparenza della raccolta dei dati dei consumatori e di fornire a quest'ultimi mezzi aggiuntivi per controllare la conservazione dei dati personali.

Sono di particolare rilevanza tre aspetti del GDPR: applicazione territoriale, consenso e responsabilità congiunta.

### 6.1.1 Applicazione territoriale

Un sito Web o un fornitore di tecnologia web situato al di fuori dell'Unione europea è soggetto al GDPR, se offre intenzionalmente beni o servizi a persone o aziende situate nell'Unione europea, e anche se controlla o prevede il comportamento, le preferenze o le attitudini personali di utenti all'interno dell'Unione Europea.

Inoltre, il GDPR protegge le persone fisiche situate nell'Unione Europea, indipendentemente dalla loro cittadinanza, residenza o altro stato giuridico.

### 6.1.2 Consenso

Se un sito Web utilizza strumenti di analisi web come cookie, script, immagini, web tracker o tecnologie per scopi pubblicitari, il sito Web deve raccogliere il consenso degli utenti.

L'opzione di consenso fornisce un mezzo semplice per proteggere la privacy: negando il consenso, un consumatore può impedire a un sito Web di raccogliere dati personali e condividerli a terzi. Se un consumatore non acconsente all'archiviazione dei dati, implica immediatamente che alcuni tipi di cookie non possono essere memorizzati, escludendo così l'intermediario.

Il GDPR non è la prima legge europea sulla privacy che ha creato un requisito di consenso relativo alle tecnologie di analisi web. Dal 2009, la direttiva sulla e-privacy (Unione Europea, 2002) aveva richiesto ai siti Web di ottenere il consenso degli utenti per l'immissione di cookie sui propri dispositivi.

### 6.1.3 Responsabilità congiunta

I siti Web utilizzano spesso servizi di terze parti come Google Analytics o Facebook Insights per il monitoraggio comportamentale e il consumo. Il GDPR riconosce questa struttura di mercato distinguendo tra coloro che determinano le finalità e le modalità del trattamento dei dati personali, e coloro che elaborano i dati personali per conto del responsabile del trattamento.

Ai sensi del GDPR, i siti Web non possono facilmente eliminare le responsabilità per le violazioni della privacy. I siti Web sono responsabili, nei confronti delle autorità sulla privacy, del trattamento congiunto dei dati con il fornitore di tecnologia Web. Devono essere, quindi, anche i siti Web, e non solo Google o Facebook, a raccogliere il consenso dell'utente per elaborare i dati personali. Inoltre, gli utenti possono anche citare in giudizio per responsabilità congiunta il sito Web per danni causati da violazioni del GDPR, anche se questi danni non sono causati direttamente dal sito web, ma da Google o Facebook.

## 6.2 Efficacia GDPR nel breve e nel lungo periodo

La protezione della privacy dei consumatori e le sue conseguenze è stata studiata da economisti, studiosi legali e scienziati informatici per diversi decenni.

Goldberg, Johnson and Shriver (2019) è il primo documento a esaminare l'impatto economico del GDPR sui siti Web europei.

Diversi altri articoli hanno studiato l'impatto del GDPR in altri settori. Jia, Jin e Wagman (2018) hanno mostrato che il GDPR ha avuto effetti negativi sugli investimenti in capitale di rischio.

Zhuo et al. (2019) hanno studiato l'impatto del GDPR sugli accordi di connessione a Internet.

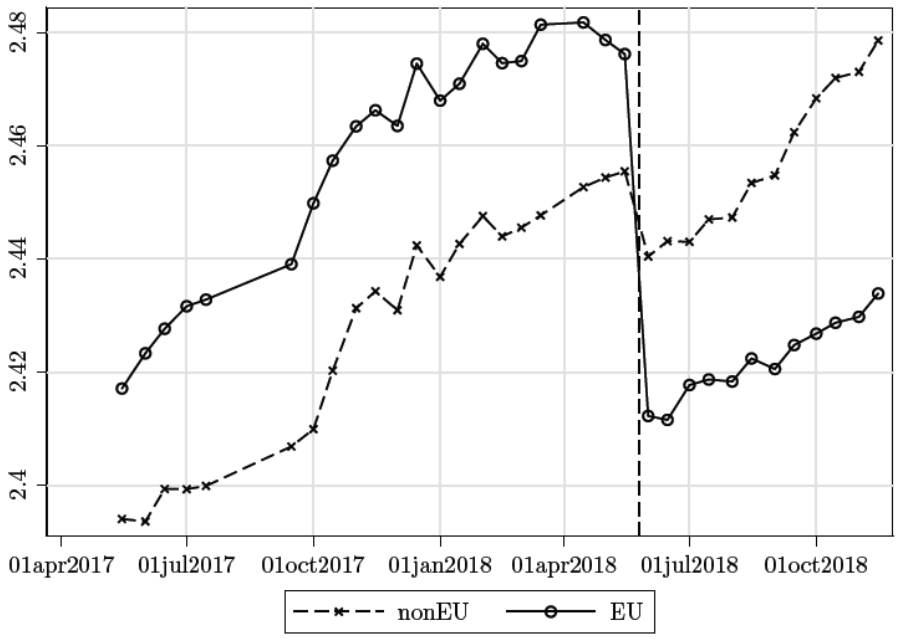
Johnson and Shriver (2019) hanno studiato i cambiamenti nella concentrazione del mercato dei fornitori di tecnologia Web a seguito del GDPR.

Il GDPR ha segnato un momento importante per la tutela dei dati personali, ma una ricerca recente ha messo in dubbio la sua efficacia e il suo impatto a lungo termine. Batikas, Peukert e altri studiosi hanno affrontato questo problema.

Utilizzando i dati di HTTPArchive, per il periodo da maggio 2017 a novembre 2018, ma studiando anche i dati a partire da maggio 2016 come controllo di robustezza, hanno definito un campione bilanciato di 110.706 siti Web che hanno osservato per oltre 33 punti nel tempo (21 punti dati prima dell'introduzione del GDPR e 12 punti dati successivi).

Una delle prime conclusioni che è emersa è che i siti Web con TLD UE si collegavano a un numero maggiore di servizi di terze parti distinti prima dell'entrata in vigore del GDPR. Si può notare, infatti, in **Figura 24**, un calo sostanziale e improvviso del numero medio di

richieste di terze parti nella settimana dopo l'introduzione del GDPR, non solo per quanto riguarda i siti Web con TLD UE, ma anche per i siti Web con altri TLD. Resta comunque più marcato il calo del numero di domini di terze parti richiesti da parte dei siti Web con TLD UE.




*"Figura 24: trend di richieste di terze parti"*

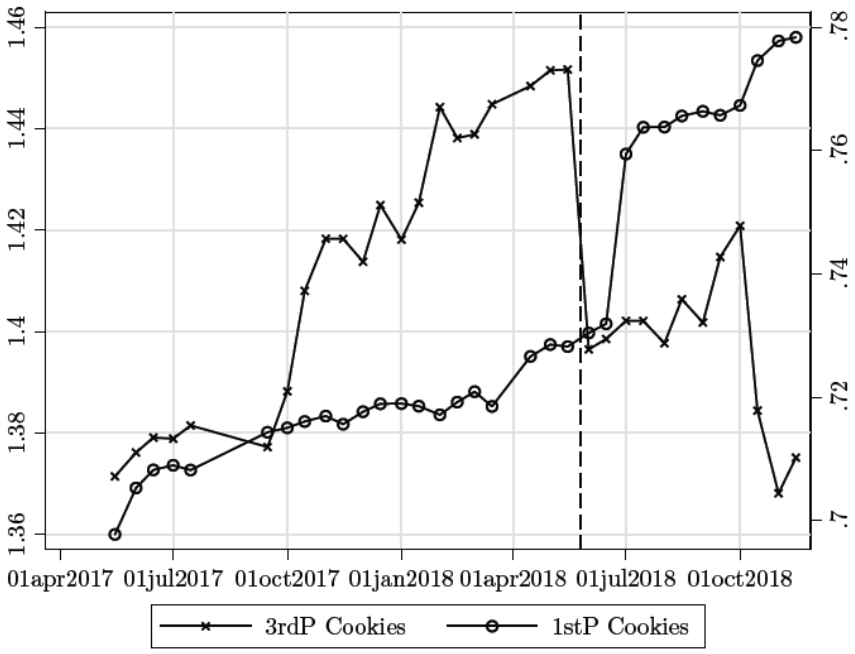
Come riportato nella **Figura 25**, inoltre, quando si controllano gli effetti fissi del sito Web e una tendenza temporale lineare, si stima che questa riduzione sia dell'8,4% (UE) e del 2,4% (non UE).

| | (1)<br>EU | (2)<br>NonEU | (3)<br>All |
|---|---|---|---|
| Trend | 0.0173*** | 0.0194*** | 0.0194*** |
| | (0.0008) | (0.0004) | (0.0004) |
| Post | -0.0840*** | -0.0245*** | -0.0245*** |
| | (0.0027) | (0.0011) | (0.0011) |
| Trend × EU | | | -0.0021** |
| | | | (0.0009) |
| Post × EU | | | -0.0596*** |
| | | | (0.0029) |
| Observations | 728112 | 2925186 | 3653298 |
| $\overline{R^2}$ | 0.8903 | 0.9033 | 0.9007 |
| Mean DV | 2.4453 | 2.4385 | 2.4399 |

*"Figura 25: effetti di richieste di terze parti"*

Analogamente, proprio come per le richieste di terzi nel complesso, vediamo una forte riduzione di cookie una settimana dopo l'introduzione del GDPR.

I cookie possono essere utilizzati per identificare lo stesso dispositivo anche se l'indirizzo IP è stato modificato, e quindi per definizione includono i dati personali.



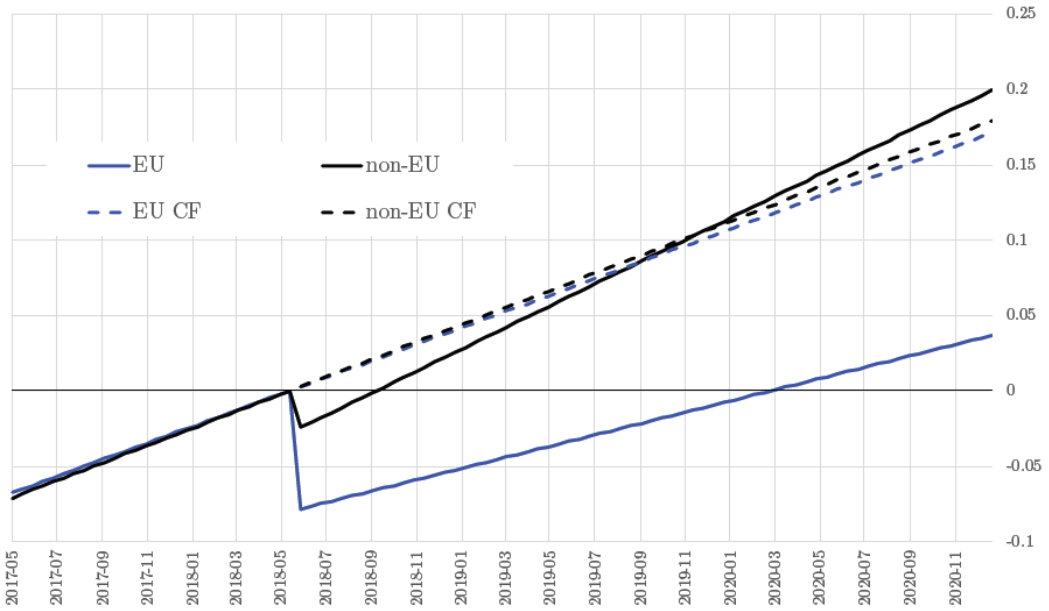

*"Figura 26: trend del numero di cookie"*

Tuttavia, solo quattro mesi dopo il GDPR, il numero di richieste di terze parti ha nuovamente raggiunto lo stesso livello direttamente prima del GDPR.

Gli effetti del GDPR di cui abbiamo discusso in precedenza sono probabilmente, quindi, di breve durata piuttosto che permanenti.

Il GDPR ha comportato una riduzione di circa il 12,5% nel totale dei cookie, ma i restanti gruppi di consumatori che non rinunciano sono tracciabili in modo più persistente.

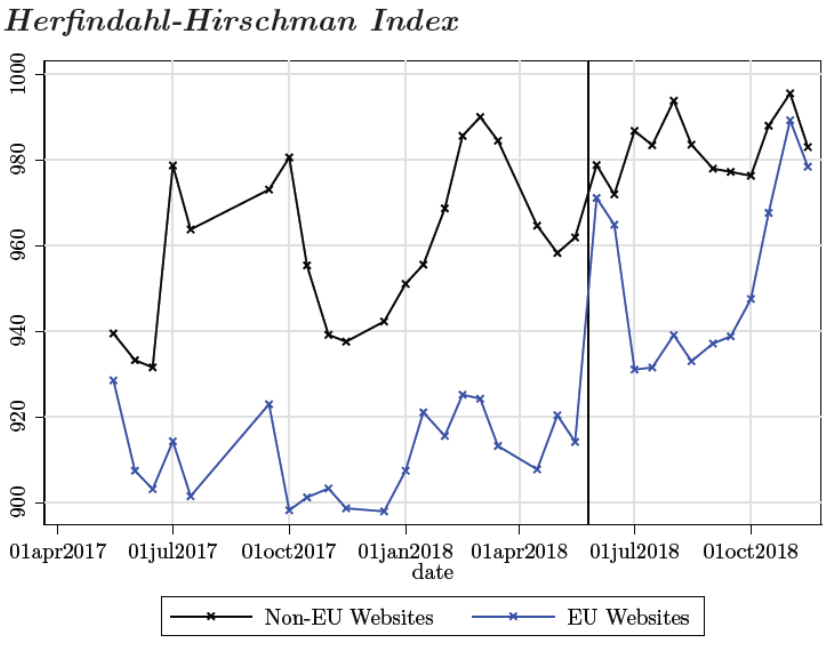
Il GDPR ha, quindi, ridotto le richieste complessive di terzi nel breve periodo, ma la tendenza crescente a lungo termine non è molto influenzata come possiamo notare nella **Figura 27**. Ciò che ne consegue è che quindi il GDPR non ha ridotto in modo efficace, nel lungo periodo, l'uso da parte dei siti Web di tecnologie di terzi che interagiscono con la privacy dei consumatori.




*"Figura 27: richieste di terze parti nel lungo periodo"*

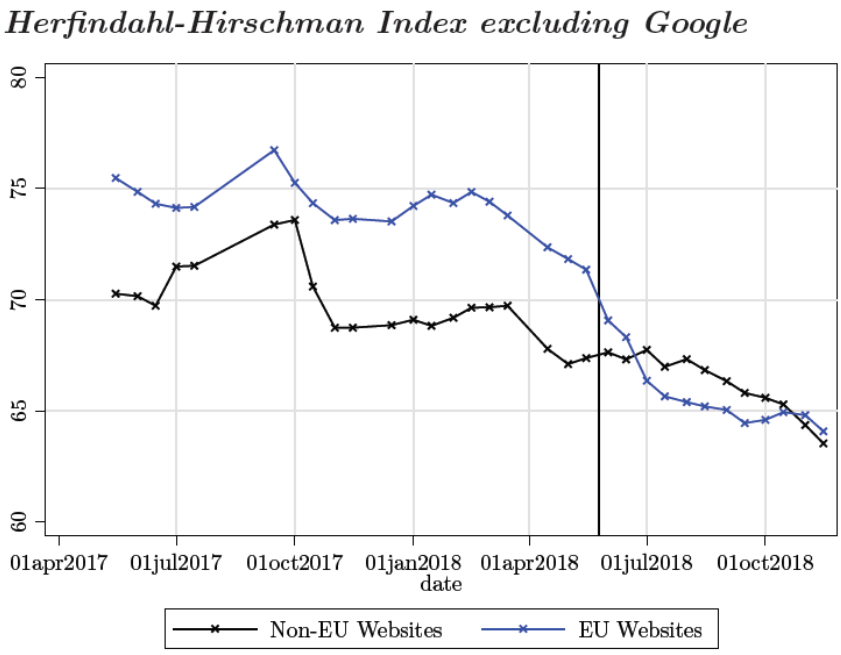
## 6.3 Conseguenze sulla struttura del mercato e sulla concorrenza.

La struttura del mercato nel mercato della raccolta dati è cambiata con l'avvento del GDPR. Nella **Figura 28**, possiamo notare come si sviluppa nel tempo l'indice Herfindahl-Hirschmann nel mercato dei siti web (somma delle quote di mercato al quadrato di tutte le aziende). L'HHI per il mercato extra-UE segue una tendenza crescente nel tempo, senza cambiamenti chiaramente visibili con l'entrata in vigore del GDPR, mentre l'HHI per il mercato UE ha subito un brusco aumento dopo l'introduzione del GDPR.




*"Figura 28: struttura del mercato dei siti web dopo il GDPR"*

Nella **Figura 29**, invece, tracciamo l'ipotetica struttura del mercato quando escludiamo l'azienda con la maggiore quota di mercato. In entrambi i mercati UE e non UE, la più grande azienda è Google. Il predominio di Google in questo mercato diventa evidente se confrontiamo la **Figura 28** con la **Figura 29**.




*"Figura 29: struttura del mercato dei siti web, senza Google, dopo il GDPR"*

L'HHI incluso Google varia all'incirca tra 900 e 1000, mentre la variazione nell'ipotetico HHI escluso Google è solo all'incirca tra 60 e 80.

Ciò suggerisce che l'aumento della concentrazione complessiva del mercato sembra essere spiegato esclusivamente da un aumento del mercato dell'impresa dominante.

Google, la tecnologia dominante in molti mercati per le tecnologie web, riesce ad aumentare la sua quota di mercato, mentre tutte le altre aziende che forniscono tecnologia web non vedono un cambiamento nella quota di mercato o subiscono perdite.

Il GDPR non ha solo portato un notevole aumento della protezione della privacy delle persone, ma ha anche aumentato considerevolmente i costi di conformità dei siti Web e dei fornitori di tecnologia web.

Normalmente le normative sulla privacy impongono alle aziende di convincere i propri consumatori a dare il proprio consenso e le aziende che offrono di più ai consumatori trovano più facile convincerle a dare il proprio consenso. Pertanto, sebbene la normativa sulla privacy imponga costi a tutti i tipi di imprese, le piccole e le nuove imprese sono

colpite in modo sproporzionato perché è più difficile per loro ottenere il consenso ai sensi del regolamento.

I consumatori hanno difficoltà a comprendere le comunicazioni sulla privacy e questa difficoltà potrebbe offrire alle grandi aziende un vantaggio rispetto alle piccole imprese, in termini di fiducia dei consumatori. Facebook, ad esempio, è considerato un servizio prezioso da molti dei suoi clienti, quindi è probabile che i consumatori acconsentano esplicitamente a dare a Facebook l'accesso ai propri dati, a differenza di un concorrente sconosciuto che non ha ancora dimostrato il suo valore.

In un mercato in cui i costi di conformità legale sono significativi, un leader di mercato come Google o Facebook potrebbe trovarsi in una migliore posizione per sostenere questi costi aggiuntivi, espandendo così la sua posizione di mercato. L'impresa dominante può, quindi, aumentare la quota di mercato e diventare più dominante.

Nel mercato dei siti Web dell'UE, Google si rafforza principalmente nei mercati in cui erano stati più forti prima del GDPR (pubblicità e analisi) con quote di mercato rispettivamente del 25,8% e del 38,4%.

La normativa sulla privacy può avere, quindi, conseguenze non intenzionali sulla struttura del mercato e sulla concorrenza. Il GDPR sembra quasi proteggere maggiormente Google che i consumatori. Può sembrare ironico, ma la fondazione della moderna legge europea sulla privacy ha aiutato i giganti della tecnologia ad aumentare le proprie quote di mercato.

# CONCLUSIONI

In questo lavoro di tesi, si è cercato di evidenziare il ruolo dei big data nel mondo odierno. Ogni volta che "clicchiamo" qualcosa sul web o effettuiamo una scelta su un dispositivo cediamo alla rete inconsapevolmente un pezzo di noi: un'opinione, un'emozione, un desiderio o un sogno. Nel mondo, in ogni minuto, sono scambiate milioni di informazioni. Ogni 60 secondi, su Facebook, sono creati 3,3 milioni di post, pubblicati 510.000 commenti e aggiornati 293.000 stati, su WhatsApp sono scambiati 29 milioni di messaggi e su Google sono effettuate 3,8 milioni di ricerche.

È evidente, quindi, come gli utenti si rapportano con un flusso continuo e ininterrotto di informazioni e come, quindi, il mondo odierno sia sommerso da un enorme quantità di dati.

I Big data sono, dunque, qualcosa che non si può più evitare. Sono ormai parte integrante della nostra vita. Sono analizzati e usati in moltissimi campi di applicazione: dalla sanità ospedaliera alla cura a distanza, dall'istruzione ai trasporti, dall'informazione giornalistica alla fruizione di contenuti multimediali, dai servizi finanziari e assicurativi alla ricerca del lavoro, dal controllo del traffico stradale alle previsioni del tempo, dall'agricoltura alla domotica e all'IOT e così via.

L'avvento di informazioni ricche e molto dettagliate sulla singola navigazione web ha iniziato a rendere possibili forme sofisticate di prezzi personalizzati, finora considerate troppo impegnative dal punto di vista informativo per essere implementate.

Inoltre, i dati sono utilizzati per targetizzare gli utenti, mostrandoli pubblicità personalizzate e aumentando la probabilità che l'advertising abbia successo.

Prezzi personalizzati e pubblicità personalizzata sono quindi due elementi fondamentali che caratterizzano la struttura dei mercati digitali. Quest'ultimi sono spesso caratterizzati da dinamiche winner take all, dove il vincitore si prende tutto il mercato. Vi è, quindi, un'elevata concentrazione di mercato, una dispersione dei profitti maggiore e forti divari economici tra le imprese.

I big data sono assimilabili a delle risorse ritenute essenziali per la realizzazione di un determinato prodotto o servizio. Ciò rende l'impresa o piattaforma che detiene quei particolari dati nella condizione privilegiata di godere di un vantaggio competitivo assoluto, che rende la sua posizione di mercato meno contendibile sia all'interno del mercato della raccolta dei dati che all'interno dell'intera filiera logistica.

Gli effetti di rete, inoltre, svolgono un ruolo fondamentale all'interno dei mercati digitali ed in particolar modo all'interno delle multi-sided platform. Sono, infatti, un'ovvia fonte di concentrazione, perché un numero maggiore di utenti migliora l'attrattiva dell'impresa dominante, portando a un numero ancora maggiore di utenti.

Dopo aver affrontato temi legati alla concentrazione nei mercati digitali, facendo attenzione in particolar modo al mercato delle App, si è analizzato il tema della privacy. Spesso, i consumatori dichiarano una forte preferenza per la privacy, ma sono disposti a rinunciare alle proprie informazioni personali anche per piccoli incentivi, presentando dei comportamenti che non si allineano agli atteggiamenti, mostrati prima di rilasciare informazioni personali, più o meno sensibili. Questa discrepanza tra atteggiamenti e comportamenti è nota come il "paradosso della privacy".

L'incertezza delle preferenze è evidente non solo negli studi che confrontano gli atteggiamenti dichiarati con i comportamenti, ma anche in quelli che stimano le valutazioni monetarie della privacy.

La ricerca corrente sta cercando di analizzare sempre più il valore che ogni singolo dato può apportare al business di questi mercati, ma un conteggio del valore effettivo dei nostri dati è molto difficile da stabilire, poiché troppi sono gli attori che beneficiano in un modo o nell'altro dei nostri dati.

Il valore, inoltre, oscilla molto a seconda del soggetto interessato: un VIP o una persona molto ricca che dichiara di possedere molti beni è più importante per le aziende e ha quindi più valore in quanto dotato di capacità di acquisto superiore, mentre i dati di una persona di ceto più basso valgono meno per il minor potere di spesa.

Nonostante queste difficoltà è stato possibile osservare il valore del dato sotto tre differenti punti di vista: disponibilità ad accettare degli incentivi monetari, disponibilità a pagare dei premi, oggettivo (algoritmo del Financial Times).

Ciò che emerge da tutti questi studi in materia è che le nostre informazioni assumono un valore maggiore quando ci viene chiesto di venderle. Siamo disposti ad accettare degli incentivi monetari purché siano alti, mentre non siamo ugualmente, con la stessa intensità, disposti a pagare dei premi per difendere le nostre informazioni sul web.

Questi due punti di vista anche se presentano dei valori molto differenti hanno in comune la soggettività nelle valutazioni monetarie della privacy.

Se ragionassimo in termini puramente economici, oggettivi, come fatto dall'algoritmo del Financial Times, il valore dei dati si abbassa ulteriormente.

Il tema della protezione della privacy e dei dati personali è apparso spesso nelle riflessioni e nei dibattiti degli ultimi anni, soprattutto nel contesto del social networking e della profilazione degli individui.

In questo contesto, nell'aprile del 2016, l'Unione Europea è stata la prima a adottare il Regolamento generale sulla protezione dei dati (GDPR), che introduce una disciplina in materia di protezione dati uniforme e valida in tutta Europa. Essendo obbligatorio in tutti i suoi elementi e direttamente applicabile in ciascuno degli Stati membri, è definito "self-executing" in quanto le sue disposizioni sono immediatamente applicabili, senza il bisogno di alcuna intermediazione.

Il GDPR ha ridotto le richieste complessive di terzi e i cookie nel breve periodo, ma la tendenza crescente a lungo termine non è molto influenzata e ciò che ne consegue è che, quindi, il GDPR non risulta essere efficace nel lungo periodo.

Il GDPR ha comportato, inoltre, conseguenze non intenzionali sulla struttura del mercato e sulla concorrenza. In un mercato in cui i costi di conformità legale sono significativi, un leader di mercato come Google o Facebook si trova in una migliore posizione per sostenere questi costi aggiuntivi, espandendo così la sua posizione di mercato.

Può sembrare ironico che la fondazione della moderna legge europea sulla privacy possa aver aiutato i giganti della tecnologia ad aumentare le proprie quote di mercato, ma è realmente ciò che è avvenuto con Google, che ha aumentato notevolmente la sua quota di mercato.

# BIBLIOGRAFIA


[1] AGCOM, Servizio economico statistico, “Big data. Interim report nell’ambito dell’indagine conoscitiva di cui alla delibera n. 217/17/CONS”, 2018.

[2] Alessandro Acquisti, Curtis Taylor, Liad Wagman, “The Economics of Privacy”, 2016.

[3] Alessandro Acquisti, Laura Brandimarte, George Loewenstein, “Privacy and human behavior in the age of information”, 2015.

[4] Alessandro Acquisti, Leslie K. John, and George Loewenstein, “What Is Privacy Worth?”,2013.

[5] Avi Goldfarb and Catherine Tucker, “Digital Economics”, 2019.

[6] Avi Goldfarb and Catherine Tucker, “Privacy Regulation and Online Advertising”, 2011.

[7] Beresford, A.R., Kübler, D.and Preibusch, S. “Unwillingness to Pay for Privacy: A Field Experiment”, WZB Working Paper, 2010.

[8] Dirk Bergemanny, Alessandro Bonattiz, Tan Ganx, “The Economics of Social Data”, 2020.

[9] E. Curry, “The Big Data Value Chain: Definitions, Concepts, and Theoretical Approaches, New Horizons for a Data-Driven Economy”, 2016.

[10] Emilio Calvano, Michele Polo, “Market Power, Competition and Innovation in digital markets: A survey”, 2019.

[11] ENISA, “Study on monetising privacy. An economic model for pricing personal information”, 2012.

[12] G. Parker e M. Van Alstyne, “Two-Sided network effects: a theory of information product design, Management Science”, 2005.

[13] Garante per la protezione dei dati personali, “Big Data e Privacy. La nuova geografia dei poteri”, 2017.

[14] Guy Aridor, Yeon-Koo Che, Tobias Salz, “Economic Consequences of Data Privacy Regulation: Empirical Evidence from GDPR”, 2020.

[15] Hann, Il-Horn, Kai-Lung Hui, Sang-Yong Tom Lee, “Overcoming Information Privacy Concerns: An Information Processing Theory”, 2007.

[16] Huberman, B.A., Adar, E. and Fine, L.R. “Valuating Privacy”, 2005.

[17] ItMedia Consulting, “L’economia dei dati. Tendenze di mercato e prospettive di policy”, 2018.

[18] J.C. Rochet e J. Tirole, “Platform competition in Two-Sided markets, Journal of the European Economic Association”, 2003.

[19] *Jentzsch, “Financial Privacy – An International Comparison of Credit Reporting Systems”, 2007*

[20] Lambrecht A., Tucker, “Can Big Data Protect a Firm from Competition?”, 2015.

[21] Lerner A. V., “The Role of Big Data in Online Platform Competition”, 2014.

[22] Mark Rysman, “The economics of Two-Sided Markets, Journal of Economic Perspectives”, 2009

[23] Michail Batikas, Stefan Bechtold, Tobias Kretschmer, and Christian Peukert, “European Privacy Laws, Global Competition and the Privacy Landscape”, 2020.

[24] Newman N., “Feature: Search, Antitrust, and the Economics of the Control of User Data”, 2014

[25] Patrick J. Kehoe, Bradley J. Larsen, Elena Pastorino, “Dynamic Competition in the Era of Big Data”, 2019.

[26] Reinhold Kesler, Michael E. Kummer, Patrick Schulte, “Competition and Privacy in Online Markets: Evidence from the Mobile App Industry”, 2019.

[27] T. Eisenmann, G. Parker e M. Van Alstyne, “Strategies for Two-Sided markets, Harvard Business Review”, 2006.

[28] Thierer A., “Relax and Learn to Love Big Data”, 2013.

# SITOGRAFIA


[1] https://managementcue.it/bigdata/9243/

[2] https://lamenteemeravigliosa.it/dati-personali-valgono/

[3] https://www.ilsole24ore.com/art/scandalizzati-facebook-vostri-dati-sono-vendita-anni-AEocgTXE

[4] https://www.devita.law/il-valore-dei-dati/

[5] https://www.garanteprivacy.it/home/docweb/-/docweb-display/docweb/3215070

[6] https://blog.osservatori.net/it_it/big-data-cosa-sono

[7] https://blog.osservatori.net/it_it/big-data-e-tutela-dei-dati-personali-%C3%A8-possibile

[8] https://www.specialistidellaprivacy.it/quanto-costano-nostri-dati-personali/#:~:text=In%20questo%20ambito%20il%20dato,utenti%20%C3%A8%20di%20%E2%82%AC%20612.

[9] https://st.ilsole24ore.com/art/tecnologie/2013-06-13/business-dati-personali-ecco-183457.shtml?uuid=AbUwxl4H

[10] https://it.wikipedia.org/wiki/Regolamento_generale_sulla_protezione_dei_dati

[11] https://www.agendadigitale.eu/cittadinanza-digitale/gdpr-tutto-cio-che-ce-da-sapere-per-essere-preparati/

[12] https://www.infodata.ilsole24ore.com/2017/05/04/quanto-vale-mercato-europeo-dei-dati/

[13] https://ig.ft.com/how-much-is-your-personal-data-worth/#axzz2WDzN1Z8V

# RINGRAZIAMENTI

*Desidero ringraziare principalmente la mia famiglia che mi ha sostenuto sia economicamente che moralmente durante tutti i cinque anni.*

*Ringrazio, quindi, i miei genitori Cosimo e Maria Rosa, che mi sono sempre stati vicini nel corso di questi anni. Ringrazio mio fratello Fabio e mia sorella Maria Chiara.*

*Un ringraziamento particolare va anche ai miei due nonni, che purtroppo la scorsa estate hanno intrapreso un nuovo viaggio. Me li immagino in cielo, contenti, a festeggiare per me. Ringrazio ovviamente anche le mie due nonne che hanno continuato ad informarsi sulla mia salute e sui miei risultati.*

*Ringrazio mio zio Giuseppe, sempre pronto a darmi consigli e tutti gli altri zii e parenti che, anche indirettamente, chiedono di me ai miei genitori.*

*Ringrazio i miei colleghi, i miei coinquilini, i miei amici e tutti coloro che mi hanno accompagnato in questo percorso.*

*In questi due anni sono cambiate tante cose. Come spesso accade, alcune persone si allontanano dalla tua vita, altri si aggiungono, altri non ti lasciano mai e altri ancora ti lasciano solo fisicamente, ma sono sempre con te per sostenerti e proteggerti.*

*Ringrazio quindi tutti coloro che davvero mi vogliono bene, tutti coloro che vogliono bene come me, senza scopo o ragione seconda.*

*Infine, ringrazio il professor Carlo Cambini che ha permesso la realizzazione di questa tesi con la sua massima disponibilità, anche a dispetto di problematiche e impedimenti vari, come la situazione Covid di quest'ultimo anno.*