

POLITECNICO DI TORINO

Corso di Laurea Magistrale in Ingegneria Gestionale



**Piattaforme digitali e siti di informazione online: studio
dell'impatto del cookie wall sull'advertising e
l'engagement**

Relatore

Prof. Carlo Cambini

Candidata

Asia Farina

Correlatore

Prof. Lorien Sabatino

Anno Accademico 2022/2023

INDICE

INTRODUZIONE	3
CAPITOLO 1 – LE PIATTAFORME DIGITALI	5
1.1 Definizione.....	5
1.2 I tipi di piattaforme online.....	7
1.3 Eliminazione o riduzione delle frizioni di mercato.....	9
1.4 Effetti di rete e struttura di prezzo non neutrale.....	9
1.4.1 Multi-homing.....	14
1.5 The “Chicken and egg” problem.....	15
1.6 Economics of data.....	17
1.6.1 I Big Data	18
1.6.2 Le caratteristiche del data market.....	21
1.6.3 Il ruolo dei dati nelle piattaforme digitali.....	25
1.6.4 Il dilemma della privacy	29
CAPITOLO 2 – IL MERCATO DEL GIORNALISMO ONLINE	34
2.1 Economia dei media.....	34
2.2 La storia del giornalismo	38
2.2.1 La prima crisi.....	42
2.2.2 La seconda crisi e la nascita del giornalismo online.....	43
2.3 I giornali online come piattaforme digitali.....	46
2.3.1 L’economia dell’attenzione.....	46
2.4 Il mercato attuale italiano	51
2.4.1 Panoramica generale	51

2.4.2 Panoramica per singola testata.....	52
2.4.3 Panoramica finanziaria del mercato dell'editoria	54
CAPITOLO 3 – ANALISI SUGLI EFFETTI DEL COOKIE WALL.....	59
3.1 L'introduzione del cookie wall.....	59
3.1.1 La posizione delle Autorità Garanti Europee	61
3.2 Introduzione all'analisi e metodologia di costruzione del dataset.....	63
3.3 Analisi preliminari.....	69
3.4 Analisi econometrica	78
3.4.1 Diff-in-Diff primo caso	80
3.4.2 Diff-in-Diff secondo caso	81
3.4.3 Diff-in-Diff terzo caso	83
3.4.4 Diff-in-Diff quarto caso	84
CONCLUSIONI.....	85
APPENDICI.....	89
BIBLIOGRAFIA	97
SITOGRAFIA.....	101

INTRODUZIONE

La presente tesi si propone di indagare gli effetti delle recenti politiche di privacy adottate da diverse testate giornalistiche e siti di informazione online italiani a partire dagli ultimi mesi del 2022. Tali strategie vincolano l'accesso ai contenuti, offrendo agli utenti due opzioni: la sottoscrizione di un abbonamento (il cosiddetto paywall) o il rilascio del consenso all'installazione di cookie e altri strumenti di tracciamento dei dati personali (il cosiddetto cookie wall). Anche la semplice visualizzazione dell'home page dei siti che adottano tale politica è così vincolata all'accettazione dei cosiddetti cookie di profilazione, utilizzati per analizzare il comportamento dell'utente e crearne un profilo dettagliato, permettendo di inviare pubblicità mirate, valutare l'efficacia delle campagne pubblicitarie e implementare strategie commerciali in linea con le preferenze dell'utente. Tale policy mira a mitigare una evidente diminuzione dei ricavi derivanti dalla vendita di pubblicità mirata, per la quale gli inserzionisti sono disposti a investire molto di più rispetto a quella rivolta ad un pubblico generico.

L'analisi si colloca nel contesto delle piattaforme digitali, approfondite nel primo capitolo, il quale ha lo scopo di fornire una comprensione delle dinamiche del mercato multi-versante e dei modelli di business che ne conseguono attraverso una revisione della letteratura, evidenziandone gli aspetti più importanti quali gli effetti di rete, la struttura di prezzo non neutrale e il ruolo dei dati, approfondendo di quest'ultimo temi come il Data Market e la privacy.

I giornali online rappresentano un tipico esempio di *multi-sided platform*, in quanto permettono l'interazione tra gli inserzionisti, i consumatori e i produttori di notizie, e generano effetti di rete indiretti, accoppiando la

domanda che i lettori hanno per gli articoli di notizie, con la domanda che gli inserzionisti hanno per un prodotto completamente diverso, la pubblicità. Nel secondo capitolo si approfondiscono, dunque, gli aspetti principali del mercato giornalistico online, partendo con una overview sull'economia dei media e l'economia dell'attenzione, fino ad arrivare ad una panoramica sul contesto attuale italiano.

Entrando nel vivo della trattazione, il terzo capitolo si focalizza sulla policy di interesse attraverso l'analisi critica dei dati raccolti e l'applicazione dei modelli di Difference-in-differences, volti ad indagare l'effetto dell'implementazione del cookie wall sull'efficacia dell'advertising online e su diverse variabili di engagement dei siti online considerati.

Infine, la conclusione dell'elaborato si prefigge di unire e dare significato alle evidenze trovate nei capitoli precedenti.

CAPITOLO 1 – LE PIATTAFORME DIGITALI

1.1 Definizione

Nella sua definizione più generica, una piattaforma è *“un mercato in cui due o più tipologie distinte di utenti (ad esempio acquirenti e venditori) possono incontrarsi per scambiare beni, servizi, informazioni”* (Martens B., 2016). Il concetto di piattaforma non è strettamente collegato al mondo digitale, ma è individuabile anche nelle strutture economiche più classiche. Basti pensare ai mercati tradizionali, in cui, grazie ad uno spazio fornito da un ente, consumatori e venditori scambiano beni e servizi, usufruendo di diversi vantaggi che altrimenti non avrebbero: ai consumatori viene offerta un’ampia varietà di prodotti, semplificando la ricerca dei beni desiderati, mentre ai produttori si permette di raggiungere uno spettro più ampio di potenziali clienti in modo più diretto. Inoltre, c’è una dinamica interdipendente tra domanda e offerta: più acquirenti attraggono più venditori sul mercato e viceversa (Martens B., 2016).

L'avvento di Internet ha reso più agevole e incentivato lo sviluppo di questo specifico modello di mercato, grazie alla disponibilità dei moderni mezzi di comunicazione che facilitano il collegamento tra i versanti e permettono di evidenziarne maggiormente i vantaggi.

Per comprendere appieno il concetto di *piattaforma digitale*, è fondamentale iniziare con un'analisi delle definizioni fornite dalle autorità di regolamentazione. In questo contesto, un importante punto di partenza è rappresentato dall’impegno della Commissione dell'Unione Europea all'interno della strategia per la definizione del cosiddetto Mercato Unico Digitale, in cui

L'ente ha attribuito alle *piattaforme digitali* un ruolo rilevante. In particolare, la Commissione UE ha definito per la prima volta la *piattaforma digitale* tramite un questionario di consultazione nel 2015 come segue: *““Piattaforma online” si riferisce a un’impresa che opera in mercati a doppio (o multi) versante, utilizzando Internet per consentire interazioni tra due o più gruppi distinti ma interdipendenti di utenti al fine di generare valore per almeno uno dei gruppi. Alcune piattaforme sono anche considerate fornitori di servizi intermediari.*

Esempi tipici includono motori di ricerca generici su Internet (come Google, Bing), strumenti di ricerca specializzati (come Google Shopping, Kelkoo, Twenga, Google Local, TripAdvisor, Yelp), elenchi di attività commerciali basati sulla posizione o alcune mappe (come Google o Bing Maps), aggregatori di notizie (come Google News), mercati online (come Amazon, eBay, Allegro, Booking.com), piattaforme audiovisive e musicali (come Deezer, Spotify, Netflix, Canal Play, Apple TV), piattaforme di condivisione video (come YouTube, Dailymotion), sistemi di pagamento (come PayPal, Apple Pay), reti sociali (come Facebook, LinkedIn, Twitter, Tuenti), negozi di app (come Apple App Store, Google Play) o piattaforme dell’economia collaborativa (come Airbnb, Uber, TaskRabbit, BlaBlaCar). I fornitori di accesso a Internet non rientrano in questa definizione”.

Con una comunicazione datata Maggio 2016, la Commissione UE ha ulteriormente affinato la sua definizione di *piattaforme digitali*, delineandone gli aspetti fondamentali:

“— possono creare e formare nuovi mercati, fare concorrenza a quelli tradizionali e organizzare nuove forme di partecipazione o di esercizio di attività economiche basate sulla raccolta, sul trattamento e sulla modifica di grandi quantità di dati;

— operano all’interno di mercati multilaterali, ma con gradi di controllo variabili sulle interazioni dirette tra gruppi di utenti;

- *beneficiano degli "effetti di rete", in virtù dei quali, generalmente, il valore del servizio aumenta con l'aumentare degli utenti;*
- *spesso si basano sulle tecnologie dell'informazione e della comunicazione per raggiungere i propri utenti in modo istantaneo e con facilità;*
- *svolgono un ruolo chiave nella creazione di valore digitale, in particolare intercettando tale valore in modo rilevante (anche attraverso l'accumulo di dati), agevolando nuove iniziative imprenditoriali e creando nuove dipendenze strategiche".*

Nella stessa sede, la Commissione ha affermato che le piattaforme digitali hanno svolto un ruolo cruciale nella società contemporanea, portando una serie di significativi vantaggi all'economia online e rivestendo un ruolo fondamentale nel creare il cosiddetto "valore digitale". Per questo motivo, le piattaforme contribuiscono al corretto funzionamento del mercato unico digitale e si distinguono per il loro contributo all'efficienza economica e per il loro ruolo centrale nell'innovazione basata sui dati. Inoltre, contribuiscono in modo sostanziale a migliorare la competitività dei mercati, estendendo notevolmente l'orizzonte delle possibilità per i consumatori e, di conseguenza, aumentandone il benessere. In ultima istanza, le piattaforme digitali potrebbero portare ad una maggiore partecipazione dei cittadini alla società e alla democrazia, semplificando l'accesso alle informazioni, specialmente tra le generazioni più giovani e oltre i confini nazionali.

1.2 I tipi di piattaforme online

Nonostante tutte le piattaforme condividano dinamiche fondamentali, possono essere implementate, organizzate e sfruttate in modo diverso dalle aziende sul mercato. Pertanto, ha senso categorizzare le piattaforme in quattro tipologie: piattaforme di transazione, piattaforme di innovazione, piattaforme integrate e piattaforme di investimento.

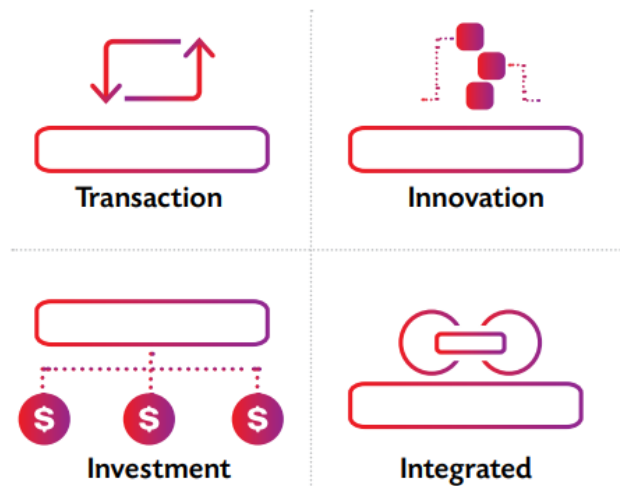


Fig. 1: Platform types. Fonte: Evans P. C. e Gawer A., 2016

- 1) *Transaction platforms*: piattaforma che agisce come intermediario, agevolando lo scambio di beni e servizi o le transazioni tra utenti, acquirenti o fornitori diversi;
- 2) *Innovation platforms*: piattaforma che funge da fondamento su cui altre aziende sviluppano tecnologie, prodotti o servizi complementari. In altre parole, è un'infrastruttura su cui altre aziende costruiscono e ampliano il proprio lavoro innovativo;
- 3) *Integrated platforms*: piattaforma che integra sia la parte di transazione che di innovazione. Questa categoria include aziende come Apple, che ha sia piattaforme di corrispondenza come l'App Store che un ampio ecosistema di sviluppatori terzi che supporta la creazione di contenuti;
- 4) *Investment platforms*: piattaforme che permettono la gestione strategica dei portafogli finanziari (Evans P. C. e Gawer A., 2016).

Il successo delle piattaforme come business model nell'economia digitale è riconducibile a specifici elementi distintivi che le caratterizzano, che permettono di creare nuovo valore per gli utenti per poi estrarlo.

1.3 Eliminazione o riduzione delle frizioni di mercato

Uno dei motivi fondamentali dietro al successo delle piattaforme è che esse *“facilitano transazioni tra diversi tipi di individui e organizzazioni che altrimenti avrebbero difficoltà a trovarsi reciprocamente”* (Evans P. C. e Gawer A., 2016). Ciò implica che questo tipo di business model sfrutta le frizioni presenti nel mercato, ovvero ostacoli che compromettono la realizzazione di una transazione economica tra le parti coinvolte e che causano un fallimento di mercato e perdita di benessere. In assenza di queste frizioni, la transazione avverrebbe in modo agevole e senza impedimenti ed è quindi possibile affermare che, grazie alle piattaforme digitali, vengono notevolmente ridotti i costi di transazione, ovvero i costi di utilizzo del mercato (Evans D. S. e Schmalensee R., 2019). Tutto questo risulta positivo sotto due punti vista: dal punto di vista di policy, la piattaforma permette di risolvere dei fallimenti di mercato, aumentare il benessere e generare quindi del valore, e dal punto di vista di business model, in quanto la piattaforma è in grado di estrarre questo valore per poi generare profitto.

1.4 Effetti di rete e struttura di prezzo non neutrale

La teoria dei mercati a due versanti è strettamente correlata alle teorie delle esternalità di rete e alla capacità di internalizzarle. Con l'espressione esternalità di rete ci si riferisce al fenomeno per cui l'utilità di un consumatore nell'utilizzare un prodotto o servizio è influenzata dal numero di utilizzatori dello stesso prodotto o servizio (He W., Li, M. e Zheng, J., 2023). Con riferimento alle piattaforme a due versanti o multi-versante, è possibile individuare due tipologie di effetti di rete:

- 1) *same side effect*: l'utilità di un utente presente sulla piattaforma è influenzata dal numero di utenti che partecipano alla piattaforma che

fanno parte dello stesso gruppo di utenti, e che si trovano quindi nel suo stesso versante;

- 2) *cross side effect*: l'utilità di un utente presente sulla piattaforma è influenzata dal numero di utenti che partecipano alla piattaforma che fanno parte dell'altro gruppo di utenti, e che si trovano quindi del versante opposto al suo.

Gli effetti di rete possono essere positivi, negativi o nulli, in base al tipo di influenza che hanno sull'utilità degli utenti.

Nella prima fase della letteratura economica sui mercati multi-versante, autori come Rochet e Tirole (2003) e Caillaud e Jullien (2003), hanno sviluppato modelli in cui si presumeva l'esistenza di effetti di rete positivi indiretti che agivano su entrambi i versanti del mercato, in entrambe le direzioni. Gli studiosi Armstrong (2006), Evans (2003), Evans & Schmalensee (2007) e Filistrucchi et al. (2013) sostengono, invece, che l'esistenza di effetti di rete indiretti in una direzione sia una condizione sufficiente per l'esistenza di un mercato multi-versante. In particolare Evans nel 2003, stabilisce le caratteristiche distintive di una piattaforma a due versanti:

- 1) la presenza di gruppi di clienti chiaramente distinti;
- 2) la capacità di un gruppo di trarre vantaggio dal coordinamento della sua domanda con il versante opposto (effetti di rete indiretti);
- 3) la possibilità di un intermediario (piattaforma) di agevolare tale coordinamento in modo più efficiente.

Gli effetti di rete cross-side possono essere rappresentati matematicamente in questo modo:

$$U_{i,A}(q) = U_{i,A}(N_A(N_B), q)$$

$$U_{i,B}(q) = U_{i,B}(N_B(N_A), q)$$

L'utilità dell'operatore economico i che si trova sul versante A dipende dal numero di utenti del versante A che a sua volta dipende dal numero di utenti sul versante B e da una certa quantità q_i e viceversa (Slide del corso di Economia e Management dei Servizi).

Il modello che segue (*Benchmark model*) dimostra che la capacità delle piattaforme di internalizzare gli effetti di rete rende le piattaforme digitali più efficienti rispetto ad imprese separate che servono i versanti in modo indipendente (Slide del corso di Economia e Management dei Servizi).

Si ipotizzi un sistema di mercato che compete sul prezzo con due versanti, con due funzioni di domanda lineari ed uguali (stessa elasticità) e costi nulli. Gli effetti di network indiretti positivi sono catturati dai parametri e_{ij} ($i=1,2; j=1,2$) che misurano di quanto aumenta la domanda sul versante j se aumenta la domanda sul versante i . Si assume che tali effetti sono positivi o nulli e che $e_{12} \times e_{21} < 1$. Si ricavano quindi le seguenti formule:

- 1) $D(p_1) = 1 - p_1$ domanda nel versante 1
- 2) $D(p_2) = 1 - p_2$ domanda nel versante 2
- 3) $q_1 = 1 + e_{21}D(p_2) - p_1$ quantità nel mercato 1 con e_{21} che misura l'effetto di rete del versante 2 sul versante 1
- 4) $q_2 = 1 + e_{12}D(p_1) - p_2$ quantità nel mercato 2 con e_{12} che misura l'effetto di rete del versante 1 sul versante 2

1° Caso: Le due imprese sono indipendenti e massimizzano il loro profitto separatamente

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

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

Massimizzando le due funzioni di profitto su p_1 e p_2 si ottengono i seguenti risultati:

$$p_1^{ind} = \frac{2 + e_{21}(1 - e_{12})}{4 - e_{12}e_{21}}$$

$$p_2^{ind} = \frac{2 + e_{12}(1 - e_{21})}{4 - e_{12}e_{21}}$$

$$\pi_1^{ind} = \frac{(2 + e_{21}(1 - e_{12}))^2}{(4 - e_{12}e_{21})^2}$$

$$\pi_2^{ind} = \frac{(2 + e_{12}(1 - e_{21}))^2}{(4 - e_{12}e_{21})^2}$$

2° Caso: un'unica piattaforma serve i due mercati e massimizza il profitto congiunto

$$\text{Max } p_1, p_2 \quad \pi = \pi_1 + \pi_2 = p_1 q_1 + p_2 q_2$$

Derivando il profitto per p_1 e per p_2 :

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

Risolvendo il sistema:

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

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

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

Se gli effetti di network tra i due versanti sono uguali ($e_{12}=e_{21}=e$), semplificando le formule dei prezzi, i prezzi risultano più bassi in entrambi i versanti nel caso di presenza di piattaforma, implicando un surplus più alto per gli utenti di

entrambi i gruppi, come anche un profitto più alto rispetto al caso di due imprese singole. Questa situazione è pareto-ottimale, ovvero è tale per cui non si può migliorare l'utilità o il benessere di un agente senza peggiorare quella degli altri. Più in generale, se gli effetti di network non sono uguali, i prezzi presi singolarmente non è detto siano inferiori rispetto al caso di imprese separate, ma la somma dei due prezzi dei versanti, e quindi anche il prezzo medio, è minore in caso di piattaforma. Anche il profitto continua ad essere più alto. Alla luce di questo modello, è possibile affermare che la piattaforma genera efficienza economica in quanto internalizza gli effetti di rete e il welfare totale aggregato è massimizzato, generando maggior valore che può essere estratto dalla piattaforma stessa sotto forma di profitto.

Rochet e Tirole, in una fase successiva del loro lavoro nel 2006, hanno rivalutato la loro precedente teoria, considerandola troppo restrittiva. Hanno quindi avanzato l'idea che, l'elemento chiave di un mercato multi-versante non risieda unicamente nella presenza di esternalità di rete, bensì nella non neutralità della struttura dei prezzi: la piattaforma può influenzare il volume delle transazioni attraverso la struttura dei prezzi, diversa per i versanti e che deve essere progettata in modo da coinvolgere entrambi i lati (Martens B., 2016). Tale tema è stato oggetto di intensa ricerca economica. Il risultato chiave è che la determinazione dei prezzi per uno dei lati del mercato non dipende solo dalla domanda e dai costi associati a quei consumatori, ma anche da come la loro partecipazione influenza la partecipazione dall'altro lato, e quindi dagli effetti di rete cross-side positivi, il profitto che ne può derivare e l'elasticità della risposta del versante opposto. In questi contesti, possono emergere situazioni apparentemente anomale, come prezzi inferiori al costo marginale o addirittura prezzi negativi, non imputabili però a pratiche anticoncorrenziali. Una piattaforma potrebbe decidere di applicare un prezzo inferiore al costo

marginale su uno dei lati del mercato se gli utenti di quel lato hanno un'alta elasticità al prezzo e la loro partecipazione attira un gran numero di partecipanti dall'altro lato, dove la domanda è relativamente poco elastica al prezzo (Rysman M., 2009). In questo modo la curva di domanda del versante meno elastico traslerebbe verso l'alto, permettendo di applicare un prezzo più elevato senza compromettere la quantità: sussidiando quindi il mercato uno, è possibile aumentare la domanda e i profitti nel mercato due più di quanto si perda nel mercato uno.

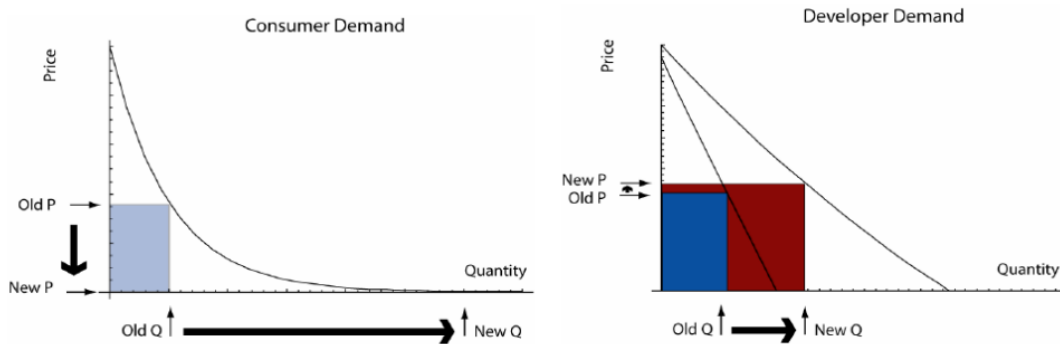


Fig. 2: Subsidization. Fonte: Slide del corso di Economia e Management dei Servizi

1.4.1 Multi-homing

L'effetto di network non si manifesta solo intra-piattaforma tra i diversi versanti, ma anche tra piattaforme diverse a causa del fenomeno del *multi-homing*. Il *multi-homing* viene definito come “la possibilità di partecipare contemporaneamente a più piattaforme” (Martens B., 2016) e può applicarsi a tutti i versanti della piattaforma o solo ad un sottoinsieme. La facilità con cui gli utenti fanno *multi-homing* dipende dagli *switching cost* tra le piattaforme e dal sistema di tariffazione adottato, che può essere basato sull'uso effettivo del servizio o prevedere tariffe fisse per l'accesso alla piattaforma. In determinati settori, gli *switching cost* sono generalmente limitati, e ciò contribuisce a mantenere un grado di monopolizzazione ridotto. Al contrario, nei social network ad

esempio, tali costi possono risultare più elevati, a causa della presenza di forti effetti di rete diretti e della complessità nell'organizzare e coordinare gruppi di utenti. La possibilità di fare *multi-homing*, in ogni caso, rende più complesso per le piattaforme trattenere gli utenti e consolidare il loro controllo sul mercato (Martens B., 2016).

Anche il *multi-homing* influenza la struttura di prezzo, infatti, nel caso in cui sul mercato siano presenti diverse piattaforme concorrenti, l'impatto della partecipazione di un lato sull'altro diventa ancora più rilevante. In presenza di due piattaforme in competizione che fissano i prezzi per i consumatori e i venditori, se la piattaforma di riferimento abbassa il prezzo per i consumatori, non solo si hanno effetti di rete positivi intra-piattaforma sul versante dei venditori, ma si attirano anche consumatori dalla piattaforma concorrente, riducendone il valore e portando ad un aumento più significativo dell'interesse dei venditori per la prima piattaforma. Di conseguenza, la formulazione dei prezzi può essere ancora più rilevante nei mercati competitivi. L'entità di questo effetto dipende in parte dal modo in cui gli agenti effettuano il passaggio da una piattaforma all'altra, ovvero se trasferiscono solo alcune delle loro transazioni o la totalità (Rysman M., 2009).

Al contrario di quello che si potrebbe pensare, nel caso in cui solo un versante della piattaforma abbia possibilità di fare *multi-homing*, in alcuni casi, può risultare più conveniente fare molta concorrenza sul versante che non fa *multi-homing* in modo da poter generare del valore e attrarre anche l'altro versante grazie agli effetti di rete.

1.5 The “Chicken and egg” problem

Stimolare l'adesione di acquirenti e venditori ad una piattaforma vuota rappresenta una sfida cruciale per le aziende che tentano di lanciare piattaforme

digitali nei mercati multi-versante. Secondo quanto previsto dalla letteratura esistente, *"nessuno si unisce finché tutti gli altri non si uniscono"*. Questo fenomeno è comunemente noto come *"The chicken and egg problem"*. La principale sfida strategica iniziale quindi, nei mercati a due versanti, è rappresentata dal superamento del cosiddetto "enigma circolare" (Spulber D. F., 2010), ovvero convincere uno dei gruppi a unirsi alla piattaforma inizialmente, prima di avvicinare l'altro gruppo attraverso l'utilizzo dei vantaggi risultanti dall'interazione e/o delle transazioni con i membri del primo gruppo. Le strategie iniziali di ingresso mirano fondamentalmente ad ottenere un numero sufficiente di partecipanti per raggiungere una "massa critica" la quale, una volta superata, permette al mercato di crescere spinto da vari meccanismi, tra cui dinamiche dipendenti dalle esternalità di rete positive (Evans D. S. e Schmalensee R., 2010). Uno studio condotto da Jørgen Veisda nel 2020, mostra come, sebbene il successo delle strategie di ingresso sia influenzato da vari aspetti legati all'organizzazione e al contesto e cambi quindi da caso a caso, gli aspetti fondamentali della gestione presentano molte similitudini. Questi aspetti riguardano la gestione delle aspettative e la governance dell'offerta. Per quanto riguarda la gestione delle aspettative, l'obiettivo della maggior parte dei manager per raggiungere il prima possibile la massa critica è quello di generare aspettative positive tra un numero significativo di individui, prima dell'effettiva operatività della piattaforma. Questo faciliterebbe infatti un'adesione di massa da parte dei potenziali partecipanti una volta che la piattaforma verrà effettivamente lanciata, risolvendo difatti il *chicken and egg problem* (Bhargava H. K., 2014). Se in alcuni casi risulta fondamentale diffondere ed evidenziare il valore qualitativo della propria proposta per alzare le aspettative dei potenziali consumatori, in mercati che presentano determinate caratteristiche contestuali legate alla richiesta di nuove soluzioni, è possibile che

la semplice introduzione di una nuova piattaforma possa essere sufficiente per un ingresso prospero. Un altro fattore cruciale per la risoluzione del *chicken and egg problem* è la gestione dell'offerta, che include il livello di accesso alle informazioni consentito, il costo di questo accesso e le regole che regolamentano l'uso della piattaforma (Gawer A., 2014). Dal momento che anche una leggera precedenza nell'attrarre i primi clienti potrebbe essere determinante, anche per un concorrente iniziale che offre un prodotto o servizio di qualità inferiore, la scelta da parte dei manager di come gestire l'offerta sulla piattaforma potrebbe essere un fattore cruciale per attrarre utenti e quindi per il suo successo finale.

Come visto, gran parte del dibattito teorico sull'efficacia delle strategie di ingresso nei mercati a due o più versanti enfatizza la rapida creazione di una vasta rete di partecipanti dai lati dell'offerta e della domanda in quanto un aumento delle dimensioni della rete aumenta il valore per tutti gli utenti al suo interno e risulta cruciale per la sopravvivenza della piattaforma (Veisdal J., 2020).

1.6 Economics of data

Nell'ambito delle loro operazioni, le piattaforme online raccolgono un vasto volume di dati riguardanti le transazioni e le informazioni personali degli utenti, noti come *Big Data*. Questi dati vengono impiegati dalle piattaforme per ottimizzare i loro servizi in base alle preferenze e ai bisogni dei consumatori. Ciò consente di migliorare i processi operativi, riducendo i costi, nonché di identificare nuove tendenze di mercato e opportunità di business. Partendo da queste informazioni e attraverso la profilazione degli utenti, le piattaforme sono in grado di anticipare e consigliare prodotti che potrebbero interessare ai consumatori. Allo stesso modo, l'uso dei dati personali degli utenti, consente ai motori di ricerca di presentare annunci pubblicitari sulle loro pagine che

potrebbero interessare all'utente, aumentando così le possibilità di interazione tra inserzionisti e consumatori. È importante notare che la capacità di offrire annunci mirati consente ai motori di ricerca di richiedere tariffe più elevate agli inserzionisti. Infine, va sottolineato che la raccolta di *Big Data* da parte degli operatori mira a migliorare i propri servizi e ad ottenere un vantaggio competitivo rispetto ai concorrenti (Cutrupi A., 2017).

1.6.1 I Big Data

Il termine *Big Data* si riferisce proprio all'enorme quantità di informazioni che ogni individuo, nel costante utilizzo dei servizi digitali, crea con le sue interazioni e azioni. La vera rivoluzione non consiste solo nel numero di dati che si producono, bensì nella rapidità con la quale gli stessi possono essere elaborati e analizzati per creare collegamenti tra essi e prevedere in che direzione potrà dirigersi la volontà di determinati individui. In questo scenario, l'aspetto quantitativo dei dati non è più il punto chiave, bensì è come le organizzazioni sfruttano le informazioni ricavate da essi a fare la differenza. Negli anni si sono 20 susseguite numerose descrizioni che sottolineano diversi aspetti caratterizzanti i big data partendo da quella di Douglas Laney nel 2001 che rappresenta una svolta sotto alcuni aspetti. L'analista di mercato articolò la definizione nota come "le tre V", secondo cui il termine Big Data descrive dati caratterizzati da:

- 1) grandi Volumi, con cui si fa riferimento alla dimensione effettiva del dataset. Le aziende sono quotidianamente inondate da un flusso di dati in costante crescita con andamento esponenziale, attraverso cui si accumulano facilmente terabyte o anche petabyte di informazioni. È la "V" a cui si faceva originariamente riferimento quando si parlava nei primi articoli di Big Data poiché all'inizio l'attenzione principale era posta quasi esclusivamente su tale caratteristica;

- 2) grande Velocità con cui si fa riferimento non solo alla rapidità con cui i dati vengono generati e acquisiti, ma anche al tempo impiegato dagli strumenti tecnologici per processare i dati. Velocizzare l'efficienza di analisi è essenziale poiché alcune attività richiedono risposte immediate (Colone B., 2020);
- 3) grande Varietà che fa riferimento alla co-esistenza di dati di natura eterogenea (provenienti da differenti fonti, umane o digitali), più o meno strutturati e organizzati in maniera e in formati differenti (Laney D., 2001).

Anche se ad oggi non esiste una definizione riconosciuta univocamente dai ricercatori e dagli esperti, la società di consulenza McKinsey&Company fornì tramite uno studio nel 2011 quella che al momento sembra essere la più quotata e completa: *“set di dati le cui dimensioni vanno oltre la capacità dei tipici strumenti software di database di acquisire, archiviare, gestire e analizzare”* (McKinsey&Co, 2011). Inoltre gli studiosi sono concordi nell'accettare un modello formato dai 3 elementi introdotti da Laney più due aggiunti di recente e chiamato delle “5V”. Oltre Volume, Velocità e Varietà ci sono:

- 1) Veridicità che indica il grado di affidabilità e attendibilità dei dati raccolti. Risulta essere un aspetto particolarmente importante per via dei cambiamenti continui a cui fonti, tecnologie e sistemi sono sottoposti ed essendo fondamentale che i dati siano consistenti. L'espressione *“Bad data is worse than no data”* indica infatti che l'analisi delle informazioni sbagliate può condurre a conclusioni fuorvianti ed errate, che risultano più dannose dell'assenza di conclusioni. Qualità e integrità sono perciò imprescindibili per produrre analisi utili e veritiere;
- 2) Valore, che rappresenta il fine ultimo nella definizione dei big data e anche della loro analisi. Questa, se effettuata in modo coerente, porta alla

creazione di un valore inestimabile. Si creano infatti conoscenze spesso valide individualmente ma essenziali vista la capacità dei dati di creare valore aggiunto attraverso le analisi predittive e le strategie innovative. Ciò accade perché le imprese che sfruttano correttamente queste elaborazioni riescono a indirizzare la gestione delle attività verso i bisogni del mercato (Colone B., 2020).

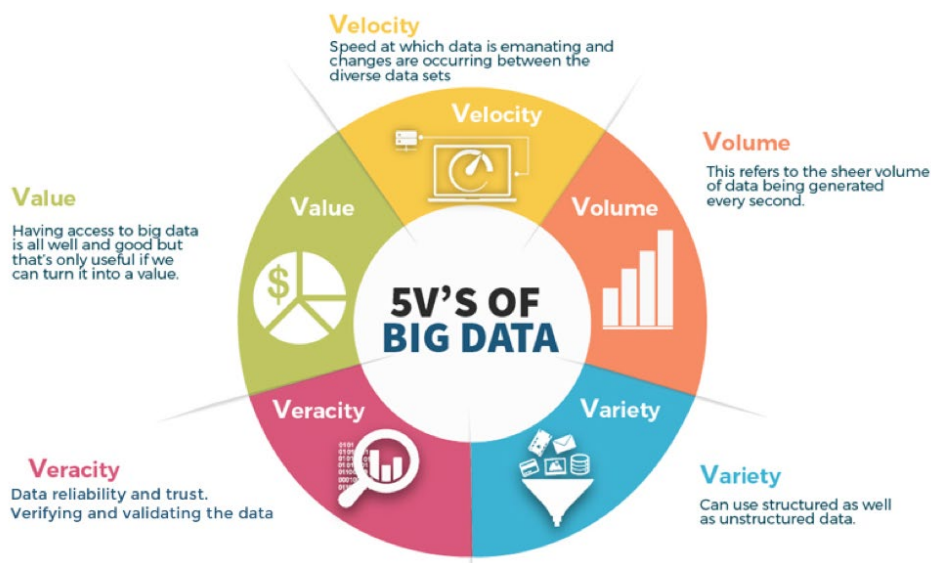


Fig. 3, Le cinque V. Fonte: www.extraordy.com

Il valore economico dei big data è stato ampiamente riconosciuto, tanto da essere definiti come “il nuovo petrolio” che alimenta l’economia (Humby C., 2006). Tale valore è intrinseco ad essi e viene estratto e valorizzato attraverso la loro “Analytics”, che consiste nell’applicazione di tecniche di analisi avanzate al fine di ottenere informazioni utili. La Big Data Analytics permette di cambiare l’approccio e aggiungere qualcosa nella gestione aziendale in diversi ambiti: si passa dall’abilità manageriale del top management, dove l’esperienza e la capacità di fare previsioni assumono un ruolo fondamentale, al decision making (processo che porta a prendere una decisione) con il supporto dei dati e della loro analisi, che diventano così i principali fattori nel prendere decisioni.

In ambito aziendale quindi, i Big Data vengono utilizzati principalmente per due ragioni: a scopo analitico, in quanto forniscono alle imprese utili intuizioni ed analisi sulle loro performance, e per offrire supporto quantitativo per il decision making. D'altro canto, i Big Data permettono alle aziende di sviluppare applicazioni e servizi in tempo reale che si avvalgono delle imponenti moli di dati digitali per creare valore per il cliente finale, valore che sarebbe impossibile ottenere altrimenti (Muhtaroglu F. C. P., Demir S., Obalı M. e Girgin, C., 2013). Il valore dei dati si sviluppa non solo da un utilizzo primario per ricavare informazioni e stilare previsioni, ma anche dai possibili utilizzi secondari ottenibili attraverso una ricombinazione di più dati. Il valore secondario è dato quindi da un riutilizzo futuro dei Big Data, al fine di ottenere informazioni estese oppure più generalizzate, a seconda della tipologia di dati uniti (Mayer-Schönberger V. e Cukier K., 2013).

1.6.2 Le caratteristiche del data market

Il dato è un pezzo di informazione rappresentato in bit e questo lo distingue dai beni e servizi ordinari. Tale caratteristica modifica i modelli economici standard, portando alla definizione di una nuova struttura di mercato. Il primo effetto rilevante in tal senso è la riduzione delle seguenti tipologie di costo:

- 1) Costi di ricerca: grazie ad Internet, è diventato notevolmente più agevole reperire e confrontare informazioni su prodotti e servizi, facilitando le transazioni online. Da un punto di vista empirico, si è constatato che i prodotti venduti online presentano un costo inferiore rispetto a quelli venduti fisicamente in negozio, in quanto, l'abbassamento dei costi di ricerca ha portato ad un aumento della concorrenza tra i rivenditori. Al contrario delle previsioni però la dispersione di prezzo tra imprese non è diminuita poiché la qualità del bene o servizio che offrono, come

l'esperienza di acquisto, spesso presentano notevoli differenze tra diversi retailer. Le aziende, inoltre, possono anche plasmare il processo di ricerca. Quando i consumatori cercano, valutano molte dimensioni di informazioni: prezzo, qualità, reputazione, spese di spedizione, tempo di consegna ecc. In presenza di costi di ricerca, intesi come tempo speso e impegno nel cercare una certa informazione, e di molte dimensioni di informazioni, le aziende possono in parte scegliere quali informazioni hanno i costi di ricerca più bassi ponendo un focus sui loro punti di forza. I bassi costi di ricerca inoltre, possono promuovere sia i prodotti superstar, ovvero quelli più inflazionati, sia quelli di nicchia, in base al tipo di algoritmo scelto per suggerire gli acquisti ai consumatori, e incrementano il matching tra venditore e consumatore riuscendo a ricavare informazioni sulla domanda di mercato;

- 2) Costi di replicazione: ciò che distingue in modo significativo i beni composti da atomi da quelli composti da bit, non è il processo di creazione, il quale risulta costoso in entrambi i casi, bensì il processo di replicazione. I bit, infatti, sono non rivali, il che significa che possono essere consumati da una persona senza ridurre la quantità o la qualità disponibile per gli altri: le informazioni possono essere condivise senza diminuire l'informazione originale. Questa caratteristica implica che, al netto di eventuali limiti legali o tecnologici, non è costoso replicare il dato una volta che è stato creato e che quindi esso può essere trasmesso ad un gran numero di persone a costi molto bassi, generando forti economie di scala: il costo di produzione dei beni e servizi digitali non cresce in modo proporzionale al numero di utenti che ne usufruiscono, ma rimane decisamente inferiore;

- 3) Costi di trasporto: al netto del costo dovuto alla congestione della rete, legato principalmente a fattori tecnologici, il costo di trasporto dei dati, risulta essere praticamente nullo o estremamente basso. Inoltre, la differenza di costo tra comunicazioni a breve o lungo raggio è trascurabile. Questo porterebbe ad un azzeramento delle distanze: individui e aziende isolate sarebbero in grado di connettersi all'economia globale e tutti i consumatori avrebbero accesso allo stesso insieme di prodotti e servizi digitali, portando ad una diffusione globale della conoscenza. Tuttavia, va notato che il luogo in cui si trova l'utente o il consumatore al momento dell'utilizzo del bene può influenzare il costo di trasporto delle informazioni. Sebbene in teoria l'utente possa accedere ai dati ovunque, le politiche adottate dagli Stati possono influire sul comportamento dei consumatori e regolare l'accesso ai dati. All'estremo, la regolamentazione può vietare determinati contenuti, rendendo l'esperienza di utilizzo di Internet diversa a seconda della località. Ciò significa che ci possono essere restrizioni su ciò che gli utenti possono accedere o vedere in base alla loro posizione geografica;
- 4) Costi di tracciamento: la digitalizzazione agevola la registrazione e l'archiviazione delle attività online degli utenti, e sia le aziende che i consumatori devono esplicitamente richiedere l'eliminazione di tali dati. Questo abbattimento dei costi di monitoraggio consente la personalizzazione e la creazione di mercati one-to-one, suscitando un rinnovato interesse per i modelli economici consolidati che considerano informazioni asimmetriche e prodotti differenziati, come la discriminazione dei prezzi e i modelli pubblicitari. Con riferimento alla discriminazione, grazie a queste dinamiche, potrebbe essere possibile applicare prezzi diversi a persone diverse grazie alla profilazione con cui

si potrebbe risalire alle *willingness to pay* dei singoli individui, riuscendo ad applicare di fatto una discriminazione di prezzo di primo grado. In questo caso i consumatori non hanno extrasurplus, estratto completamente dai produttori, ma tutti i consumatori hanno accesso al mercato. Per quanto riguarda l'advertising, grazie al tracciamento, è possibile fare una pubblicità mirata per il singolo consumatore e rilevare l'impatto e l'efficacia delle campagne;

- 5) Costi di verifica: la verifica dell'identità e della reputazione del dato comporta determinati oneri. Un esempio di modello che si basa su questo tipo di verifica è rappresentato dalle recensioni online, che gli utenti forniscono senza alcun costo. Tale modello consente a tutti gli utenti di ridurre a zero il proprio costo per la convalida delle informazioni. Inoltre, nel caso di recensioni positive, permette al fornitore del servizio di incrementare i propri profitti e ricavi. Tuttavia, è importante notare che questo modello presenta un limite: non tutti i consumatori utilizzano gli strumenti di recensione e, nel caso lo facciano, non esiste una garanzia che le recensioni siano veritiere (Goldfarb A. e Tucker C., 2019).

Le caratteristiche uniche dei dati come asset economici, inoltre, fanno sì che il valore che un'azienda attribuisce ad un dato dipenda dalla sua capacità di combinare quel dato con quelli già in suo possesso. Maggiore è la complementarità, maggiore sarà il prezzo che l'azienda sarà disposta a pagare. Questo significa che non esiste un mercato unico che stabilisca un prezzo standard per i dati: ogni azienda valuta in base alle proprie esigenze e opportunità. Inoltre, lo scambio di dati avviene attraverso contrattazioni private tra gli utenti, aggiungendo un elemento di personalizzazione e negoziazione nel processo. Un'altra caratteristica importante è che il valore dei dati dipende

dall'accuratezza delle previsioni che possono essere generate utilizzandoli. La maggior parte delle volte, infatti, essi sono impiegati per scopi predittivi, come le preferenze degli utenti su prodotti e servizi o previsioni di acquisti futuri. Infine, l'autenticità e l'utilità dei dati sono difficili da stabilire in anticipo e diventano evidenti solo quando vengono effettivamente utilizzati. È complesso determinare in anticipo se un dato contribuirà o meno a migliorare le prestazioni delle previsioni desiderate. Pertanto, la valutazione del valore di questo asset richiede una previsione informata su come esso influenzerà il processo predittivo, rendendo il processo di valutazione ancora più sofisticato (Slide del corso di Economia e Management dei Servizi).

1.6.3 Il ruolo dei dati nelle piattaforme digitali

Le piattaforme digitali acquisiscono i dati attraverso tre canali principali. In primo luogo, vi sono dati forniti volontariamente, ossia informazioni che gli utenti forniscono intenzionalmente mentre utilizzano un prodotto o servizio. Questa categoria comprende elementi come nomi, indirizzi email, immagini o video, informazioni di calendario, ma anche dati più strutturati come recensioni o post sui social media. In seguito, parte dei dati viene osservata. Ad oggi, molte attività lasciano una traccia digitale, e per "dati osservati" si intendono informazioni comportamentali raccolte automaticamente dalla piattaforma grazie alle azioni di un utente sul web. Ad esempio, il movimento delle persone o i percorsi di un veicolo possono essere tracciati tramite i cellulari. In generale, ogni click su una pagina web può essere memorizzato dal sito web e software di terze parti monitorano il modo in cui i visitatori si comportano: per questo motivo lo sviluppo di Internet e delle attività digitali fa sì che vengano generate enormi quantità di dati. Per concludere, ci sono i dati inferiti, ossia ottenuti tramite una trasformazione non elementare di dati volontari e/o osservati, pur mantenendo una correlazione con un individuo. Ciò può includere la creazione

di determinati profili di acquirenti, ottenuti da algoritmi di raggruppamento o predizioni sulla propensione di una persona all'acquisto di un prodotto o sulla sua valutazione del credito.

Un altro aspetto da valutare sono le diverse forme di utilizzo del dato. Sotto questo punto di vista sono quattro le categorie più rilevanti: l'uso non anonimo dei dati a livello individuale, l'uso anonimo dei dati a livello individuale, i dati aggregati e i dati contestuali. La prima categoria, comprende tutte le forme di utilizzo del dato individuale, sia volontario, osservato o inferito, con lo scopo di fornire un servizio all'individuo che li ha forniti, come ad esempio, un'applicazione musicale che utilizza i dati sulle canzoni ascoltate da un utente per offrire raccomandazioni su nuovi artisti che potrebbero piacergli. La categoria che si concentra sull'uso anonimo dei dati a livello individuale, invece, comprende tutti i casi in cui i dati a livello individuale vengono trattati in modo anonimo. Pur essendo necessario l'accesso ai dati a livello individuale, l'obiettivo non è fornire un servizio diretto alla persona che ha generato i dati inizialmente, ma, ad esempio, addestrare algoritmi di apprendimento automatico utilizzati per scopi non strettamente legati agli scopi originali per cui i dati sono stati raccolti. In specifici contesti, le informazioni ottenute potrebbero essere utilizzate per offrire un servizio migliorato a parte degli individui che hanno contribuito con i dati, ma comunque non in modo singolo e puntuale. I dati aggregati fanno un uso combinatorio dei dati generati dagli utenti, aggregandoli in modo definitivo per ottenere ad esempio informazioni sulle vendite, statistiche nazionali e i bilanci delle aziende. Infine, i dati contestuali riguardano tipicamente informazioni su determinate circostanze, come dettagli sulla rete stradale, dati satellitari e dati di mappatura (Crémer J., de Montjoye Y. A. e Schweitzer H., 2019).

Una delle tecnologie più diffuse che facilita la raccolta e la rivendita di dati individuali è basata sui cookie e su mezzi correlati per registrare i dati di navigazione. I cookie sono piccoli file che un sito web inserisce nel browser di un utente per registrare informazioni sulla sua visita. I fornitori di dati utilizzano vari siti web partner per inserire cookie nei computer degli utenti e raccogliere dati. In particolare, la prima volta che un utente visita un sito partner (come ad esempio un sito di viaggi), un cookie viene inviato al suo browser, registrando ogni azione intrapresa durante quella sessione di navigazione (come le ricerche di voli). Se lo stesso utente visita un altro sito web partner (come un rivenditore online), le informazioni contenute nel suo cookie vengono aggiornate per riflettere la cronologia di navigazione più recente. Il fornitore di dati mantiene così un profilo dettagliato e sempre aggiornato per ciascun utente, e elabora segmenti di caratteristiche dei consumatori, basandosi sul comportamento di navigazione di ogni individuo (Bergemann D. e Bonatti A., 2015).



Fig. 4, I cookies. Fonte: www.wired.it

I dati sono fonte di guadagno per le piattaforme digitali, grazie al loro ruolo cruciale nell'online advertising. Grazie all'ambiente online e ai dati, si aprono

due possibilità fondamentali per gli advertisers: la "*targetability*", cioè la capacità di indirizzare la pubblicità verso gli utenti che sono probabilmente più interessati a quel prodotto o servizio, e la "*measurability*", che consiste nella valutazione dell'efficacia della campagna pubblicitaria.

Il "targeting" rappresenta la possibilità per un inserzionista di mostrare un annuncio specifico ad un gruppo selezionato di potenziali acquirenti anziché a tutti gli utenti. Questo è un vantaggio significativo rispetto alle piattaforme offline, dove la personalizzazione è limitata. Le piattaforme online consentono un targeting preciso basato sui dati utente, riducendo i costi di ricerca che l'inserzionista dovrebbe sostenere per identificare i potenziali consumatori, un servizio che invece viene fornito dalla piattaforma stessa. Questa efficienza è resa possibile dalla capacità dei sistemi digitali di gestire e archiviare grandi quantità di dati ad un costo molto basso, talvolta addirittura nullo per la piattaforma. Inoltre, l'uso diffuso dei dispositivi digitali (smartphone, pc, ecc.) permette il targeting mirato riuscendo a ricollegare i dati ad un certo individuo. Attraverso l'analisi di questa attività online, la piattaforma può approssimare le preferenze e i bisogni dell'utente, grazie alla cronologia di navigazione la quale fornisce informazioni sia sulle preferenze a breve termine (ad esempio, una recente ricerca di un determinato prodotto), sia sulle preferenze generali (come il tipo di musica preferito o il tipo di abbigliamento). Dal punto di vista economico, si ha un gruppo di potenziali consumatori di cui non sono note le preferenze, e un gruppo di inserzionisti desiderosi di offrire prodotti più vicini alle esigenze del cliente. La piattaforma può quindi inviare un annuncio mirato, che aumenta il suo valore agli occhi degli inserzionisti, già determinato dal numero di utenti raggiungibili grazie agli effetti di rete, ottenendo così un livello più alto di profitto.

La misurabilità, invece, è la capacità di valutare l'efficacia di una campagna pubblicitaria in modo da poterla ottimizzare. Questo è possibile grazie alle piattaforme, sia perché consentono alle aziende di collegare la visualizzazione dell'annuncio con il comportamento futuro dell'utente attraverso la sua cronologia di navigazione, e sia perché consentono di condurre con facilità esperimenti sul campo, in cui a un gruppo casuale di utenti viene mostrato un certo tipo di annuncio per valutarne l'efficacia in modo preventivo (Slide del corso di Economia e Management dei Servizi).

1.6.4 Il dilemma della privacy

Come visto, i nuovi modelli di business digitali hanno portato ad un notevole ampliamento e affinamento nella raccolta dei dati dei consumatori da parte delle imprese. Questi avanzamenti sono stati seguiti da sviluppi nell'ambito del *machine learning* e di altre tecnologie di elaborazione dati, che hanno consentito alle aziende di trasformare tali dati in prodotti e servizi di successo, ottenendo considerevoli profitti ad esempio attraverso l'advertising. Tuttavia, parallelamente a ciò, si è assistito ad un aumento dei casi di violazione della privacy di risonanza mediatica e ad una crescente sensazione di impotenza tra i consumatori, i quali spesso si sentono privi di controllo sui loro dati (Aridor G., Che Y. K. e Salz T., 2020). In particolare, i dati raccolti presentano i seguenti rischi in ambito privacy:

- 1) *data persistence*: una volta che un dato è stato creato, può rimanere accessibile per molto più tempo rispetto alla vita dell'individuo che lo ha generato, poiché i costi di archiviazione sono molto bassi;
- 2) *data repurposing*: non si hanno certezze su come la piattaforma potrebbe utilizzare il dato in futuro, se sarà condiviso o replicato e quanti/quali utilizzi potrebbe avere;

3) *data spillovers*: i dati possono generare informazioni anche su utenti che non hanno direttamente utilizzato l'applicazione. Ad esempio, se un utente sceglie di non essere tracciato, ci sono altri utenti simili che potrebbero aver acconsentito e che aiutano la piattaforma a profilare anche l'utente che ha optato per la non tracciabilità (Slide del corso di Economia e Management dei Servizi).

La storia della Protezione dei Dati in Europa è stata plasmata da una serie di eventi significativi che hanno marcato importanti pietre miliari nel corso degli anni. Nel 1948, con l'adozione della Dichiarazione Universale dei Diritti dell'Uomo, è stato incluso per la prima volta il diritto alla privacy e nel 1981 il Consiglio d'Europa adottò la "*Data Protection Convention*", dando ancora più risalto a livello legale a tale diritto. Il 1995 segnò un momento epocale con l'introduzione della prima direttiva europea sulla Protezione dei Dati, che pose le basi per un quadro normativo unificato all'interno dell'Unione Europea. Successivamente, nel 2002, venne introdotta la "Direttiva sulla Privacy e le Comunicazioni Elettroniche", che affrontò specificamente la protezione dei dati nell'ambito delle comunicazioni elettroniche, includendo disposizioni sui cookie elettronici. Un altro passo importante fu fatto nel 2012, in cui è stata rilasciata la bozza di proposta per il GDPR e, solo nel 2016 dopo quattro anni di discussione, il Parlamento Europeo votò a favore dell'approvazione del regolamento, portando ad una riforma significativa della normativa sulla privacy. Infine, nel 2018, il GDPR entrò in vigore definitivamente, confermando l'impegno dell'Unione Europea per la protezione dei dati personali.

DATA PROTECTION

A brief timeline

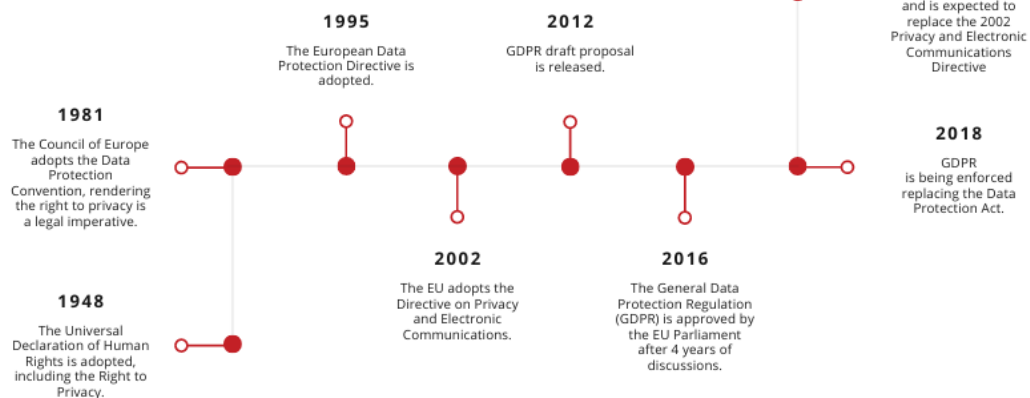


Fig. 5, Timeline della Data Protection in Europa. Fonte: www.avangarde-software.com

In particolare, il *General Data Protection Regulation* (GDPR) dell'Unione Europea è stato applicato in tutti gli Stati membri a partire dal 25 maggio 2018, coprendo i dati personali di tutti i residenti dell'UE, indipendentemente dalla località del trattamento. Esso rappresenta una pietra miliare nell'evoluzione del quadro normativo europeo sulla privacy e non solo, vista l'ampia influenza a livello globale (Goddard M., 2017). Il regolamento fornisce la seguente definizione di dato personale: *“qualsiasi informazione riguardante una persona fisica identificata o identificabile («interessato»); si considera identificabile la persona fisica che può essere identificata, direttamente o indirettamente, con particolare riferimento a un identificativo come il nome, un numero di identificazione, dati relativi all'ubicazione, un identificativo online o a uno o più elementi caratteristici della sua identità fisica, fisiologica, genetica, psichica, economica, culturale o sociale”* e definisce il suo trattamento come segue: *“qualsiasi operazione o insieme di operazioni, compiute con o senza l'ausilio di processi automatizzati e applicate a dati personali o insiemi di dati personali, come la raccolta, la registrazione, l'organizzazione, la strutturazione, la conservazione, l'adattamento o la modifica, l'estrazione, la consultazione, l'uso, la comunicazione mediante trasmissione, diffusione o qualsiasi altra forma di messa a*

disposizione, il raffronto o l'interconnessione, la limitazione, la cancellazione o la distruzione". Il GDPR delinea inoltre i principi che devono essere applicati al trattamento:

- 1) *liceità, correttezza e trasparenza: "i dati sono trattati in modo lecito, corretto e trasparente nei confronti dell'interessato". Il trattamento deve essere consentito dalla legge (quindi su basi legali indicate dal GDPR), per le sole finalità comunicate agli interessati e l'informativa privacy deve essere resa con linguaggio semplice e chiaro, completa ed intellegibile;*
- 2) *limitazione delle finalità: "i dati sono raccolti per finalità determinate, esplicite e legittime, e successivamente trattati in modo che non sia incompatibile con tali finalità";*
- 3) *minimizzazione dei dati: "i dati sono adeguati, pertinenti e limitati a quanto necessario rispetto alle finalità per le quali sono trattati";*
- 4) *esattezza: "i dati sono esatti e, se necessario, aggiornati; devono essere adottate tutte le misure ragionevoli per cancellare o rettificare tempestivamente i dati inesatti rispetto alle finalità per le quali sono trattati";*
- 5) *limitazione della conservazione: "i dati sono conservati in una forma che consenta l'identificazione degli interessati per un arco di tempo non superiore al conseguimento delle finalità per le quali sono trattati; i dati personali possono essere conservati per periodi più lunghi a condizione che siano trattati esclusivamente a fini di archiviazione nel pubblico interesse, di ricerca scientifica o storica o a fini statistici";*
- 6) *integrità e riservatezza: "i dati sono trattati in maniera da garantire un'adeguata sicurezza dei dati personali, compresa la protezione, mediante misure tecniche e organizzative adeguate, da trattamenti non autorizzati o illeciti e dalla perdita, dalla distruzione o dal danno accidentali";*

- 7) responsabilizzazione: *“il titolare del trattamento è competente per il rispetto dei principi sopra elencati e in grado di provarlo”*.

Aridor et al. in uno studio del 2020 hanno cercato di individuare gli effetti dell'applicazione del GDPR sulla capacità delle aziende di raccogliere dati dei consumatori, prevedere il loro comportamento e monetizzare tramite la pubblicità online. In particolare, la loro ricerca ha evidenziato una diminuzione del 12,5% del numero di consumatori tracciati dovuta alla possibilità di negare il consenso per tale pratica ma, al tempo stesso, la *browsing history* degli utenti ancora osservabili risulta più lunga e, di conseguenza, di migliore qualità. Tali utenti risultano inoltre più prevedibili. Coerentemente con questa ipotesi, si osserva che il valore medio dei restanti consumatori per gli inserzionisti è aumentato, compensando la maggior parte delle perdite dovute alla sottrazione degli utenti che non prestano il consenso. Infine, è riscontrato che la capacità di prevedere il comportamento dei consumatori non peggiora significativamente. Tutto questo potrebbe significare che le imprese non vengano danneggiate in modo significativo dal GDPR, che in ogni caso sicuramente favorisce i consumatori che decidono di non prestare il proprio consenso, proteggendo la loro privacy. Per quanto riguarda i consumatori che scelgono di farsi tracciare, il loro benessere dipende da come le aziende utilizzano i loro dati. Se i dati vengono impiegati per personalizzare la pubblicità e i servizi in base alle loro esigenze, potrebbero anch'essi beneficiare delle leggi sulla privacy, considerando anche tutti gli obblighi a cui le aziende sono tenute grazie al GDPR. Tuttavia, se i dati vengono utilizzati per ottenere un surplus dai consumatori, ad esempio applicando una discriminazione di prezzo di primo grado, il consenso prestato potrebbe recare loro danno.

CAPITOLO 2 – IL MERCATO DEL GIORNALISMO ONLINE

2.1 Economia dei media

L'economia dei media comprende un'ampia varietà di attività focalizzate sulla creazione, diffusione e commercializzazione di contenuti editoriali, che possono assumere diverse forme, con una predominanza di testi scritti, contenuti audio e video. Questa industria ad oggi si trova in stretta sinergia con altri due settori, l'informatica e le telecomunicazioni, identificando un unico macro-settore economico e industriale della comunicazione. Nonostante abbiano origini e sviluppi diversi, queste tre componenti hanno svolto un ruolo cruciale nella storia dell'industria dei media. In particolare, l'editoria coinvolge le attività di creazione di contenuti di comunicazione come giornali, libri, dischi, film e altri prodotti audiovisivi ed ha preso il volo nel XV secolo, grazie all'invenzione della tecnica di stampa di Johann Gutenberg. Le telecomunicazioni, includono le attività di fornitura di infrastrutture e servizi di trasmissione dei contenuti e hanno preso piede nella prima metà del XIX secolo con l'introduzione del servizio telegrafico basato sulla tecnologia di Samuel Morse. Infine, l'industria informatica ha iniziato a prosperare nella seconda metà del XX secolo, sviluppando applicazioni militari sulla base della tecnologia di John V. Atanasoff e Clifford Berry e copre le attività di aziende che forniscono hardware e software per il trattamento automatico delle informazioni e le loro applicazioni nel contesto della produzione, distribuzione e consumo di contenuti audiovisivi e testuali.

L'industria dei media rappresenta dunque un settore diversificato, composto da elementi distinti che possono essere gestiti da differenti entità economiche, ma che operano in modo sinergico per offrire un servizio unificato agli utenti finali. Una parte cruciale di questa industria è dedicata alla creazione di materiale editoriale, il “contenuto”, che spazia in termini di espressività, complessità intellettuale e finalità, e va dall'informazione all'educazione all'intrattenimento e ed è influenzato da valori economici e culturali, mentre l'altra parte, definita anche come “contenitore”, è composta dall'infrastruttura materiale che facilita la distribuzione e la trasmissione di tali contenuti nel tempo e nello spazio, la quale segue le dinamiche tipiche della produzione industriale, con economie di scala, miglioramenti di produttività e un costante impatto dell'innovazione tecnologica.

Questa industria è caratterizzata da alcuni elementi distintivi rispetto ad altri settori economici che rendono i media entità uniche che comportano decisioni specifiche in termini di produzione, distribuzione e consumo. Queste peculiarità riguardano diversi aspetti dell'attività economica:

- 1) beni pubblici e beni privati: nel settore dei media, come visto, è possibile distinguere tra contenuti e contenitori. I contenuti sono immateriali e rientrano nella categoria dei "beni pubblici" in quanto sono caratterizzati dalla non rivalità (il consumo da parte di un individuo non riduce la possibilità di consumo da parte di altri) e dalla non escludibilità del consumo (è impossibile o difficile escludere qualcuno dal loro utilizzo). I contenitori, invece, sono tangibili e rientrano nella categoria dei "beni privati". Questi beni possono essere esclusi da altri individui, ad esempio, nel caso di un film proiettato in un cinema o di servizi a pagamento. I contenitori materiali, inoltre, si deteriorano con l'uso

ripetuto e col tempo, mentre il valore dei contenuti medialti non si esaurisce con il consumo, motivo per il quale possono essere riproposti ad un numero sempre crescente di persone;

- 2) doppio valore del prodotto: l'industria dei media si distingue per la produzione di contenuti che possiedono due forme di valore distinti: il valore economico, legato al costo di produzione e alle vendite, e il valore culturale, che riguarda gli aspetti formali e contenutistici e può essere valutato solo qualitativamente. Questi due valori influenzano il mercato dei contenuti e, a volte, in modo non uniforme. Ci sono casi in cui opere di successo commerciale possono essere considerate di scarso valore culturale e viceversa. Le politiche e gli interventi istituzionali nel settore dei media prendono spesso in considerazione entrambi i tipi di valore, adottando strategie diverse per proteggerli e sostenerli;
- 3) struttura di costo: un altro aspetto distintivo sono i costi di produzione, che si compongono di costi fissi (indipendenti dalla quantità prodotta) e costi variabili (che variano in base alla quantità prodotta). In questo settore, i costi fissi sono notevolmente elevati, mentre i costi variabili sono minimi o addirittura nulli. Tale circostanza implica che all'aumentare del numero di replicazioni di un contenuto multimediale, il costo per ciascuna duplicazione si riduce poiché le spese fisse vengono ripartite su un maggior numero di unità. La fase più impegnativa e cruciale è la produzione del prototipo o "prima copia", che assorbe gran parte delle risorse produttive strategiche e definisce le caratteristiche fondamentali del prodotto. La distribuzione del contenuto richiede un ulteriore investimento, il cui ammontare può variare in base al tipo di supporto utilizzato. Nel complesso, i costi sostenuti da un'impresa mediale per fornire ulteriori contenuti editoriali si riducono

progressivamente, incrementando di conseguenza il margine di guadagno;

- 4) modalità di finanziamento: nel contesto dell'industria dei media, emergono due tipi di prodotti, destinati ai due versanti del *two-sided market*. Il primo riguarda la creazione e la distribuzione di contenuti, mentre il secondo si concentra sulla formazione di un pubblico, attirato da tali contenuti. In alcuni casi, entrambi i prodotti generano entrate economiche. Ad esempio, un editore di giornali vende il proprio prodotto ai lettori e offre agli inserzionisti pubblicitari l'opportunità di raggiungere questo pubblico. Queste risorse economiche possono combinarsi in vari modi, a seconda del modello di business adottato. L'incremento degli spazi pubblicitari su un giornale può portare ad una diminuzione della soddisfazione dei lettori e quindi ad una riduzione delle vendite. Tuttavia, questa riduzione può essere compensata dalla riduzione del prezzo del giornale, sussidiata dagli introiti pubblicitari maggiori. In generale, all'aumentare dell'esposizione pubblicitaria, diminuisce il prezzo di acquisto o di accesso, e viceversa. Questo spettro va dall'accesso gratuito ai contenuti mediali finanziati interamente dalla pubblicità, all'accesso a pagamento in media specializzati dove l'advertising ha una presenza marginale o nulla. Inoltre, soprattutto in Europa, si aggiunge un'altra modalità di finanziamento rappresentata dai contributi pubblici, che variano a seconda dei media e delle istituzioni coinvolte;
- 5) ciclo di vita dei prodotti: i prodotti mediatici possono essere classificati in due ampie categorie in base alla durata della loro rilevanza economica. Una prima categoria riguarda i "prodotti di persistenza," noti anche come "oggetti di utilità ripetuta," che possono essere riutilizzati nel

tempo, continuando a generare ricavi. Questa classe comprende opere come film, romanzi, brani musicali e altri contenuti la cui domanda può perdurare per un lungo periodo. Gli autori, editori o detentori dei diritti possono sfruttare in modo esclusivo tali contenuti per un periodo stabilito per legge, ma il loro valore economico persiste finché esiste un pubblico disposto a investire per fruirne. Esempi includono opere della classicità greca e latina, così come produzioni di epoche passate, ancora oggi pubblicate e vendute. La seconda categoria è quella dei "prodotti in scorrimento", i quali perdono rapidamente interesse nel pubblico e il loro valore commerciale si esaurisce quasi immediatamente dopo la distribuzione, o in un breve lasso di tempo. Questo gruppo abbraccia giornali, notiziari radiofonici e televisivi, talk show, programmi a quiz, la maggior parte delle trasmissioni sportive in diretta e tutti quei programmi che non hanno la possibilità di essere riproposti in radio o in televisione. Anche quotidiani e periodici appartengono a questa categoria, poiché la loro rilevanza economica solitamente dura solo un giorno, una settimana o un mese. Il ciclo di vita di un prodotto determina sia le peculiarità della sua rete distributiva sia le strategie di marketing associate ai suoi contenuti. I prodotti in scorrimento, caratterizzati da una rilevanza economica molto breve, necessitano di una rete di distribuzione che li renda accessibili al più ampio pubblico possibile nel minor tempo possibile, mentre i distributori dei prodotti persistenti offrono condizioni di accesso prolungate, in corrispondenza di un ciclo di vita economica del prodotto più esteso (Richeri G., 2012).

2.2 La storia del giornalismo

Il giornalismo rappresenta uno dei pilastri fondamentali nel panorama mediatico mondiale: fin dai primordi, i giornali hanno giocato un ruolo

determinante nella formazione dell'opinione pubblica, esercitando un'influenza significativa su discussioni politiche, sociali e culturali.

Le prime forme organizzate di comunicazione emersero nell'antica Roma, con gli "*Annales maximi*", gli "*Acta diurna*" e gli "*Acta senatus*", che diffondevano accadimenti importanti e atti ufficiali e, nel Duecento, nacquero i primi servizi postali che utilizzavano corrieri per trasmettere lettere di avviso o novelle. Queste comunicazioni erano principalmente rivolte a persone di un alto ceto sociale come uomini d'affari, ma contenevano spesso informazioni di vario tipo.

L'invenzione della stampa da parte di Johann Gutenberg nel 15° secolo rappresentò la prima svolta epocale, grazie alla quale nel Cinquecento, sorsero numerose tipografie nelle principali città europee, le quali divennero centri culturali di rilievo, rivoluzionando la circolazione delle notizie tramite bollettini commerciali e ufficiali che assunsero invece la forma di libri di notizie, per poi trasformarsi in fogli numerati che divennero di dominio pubblico. Questi strumenti furono i precursori dei primi periodici ad ampia diffusione, le cosiddette "gazzette", pubblicazioni principalmente a due o quattro pagine di piccolo formato, che presentavano resoconti sulle corti e dettagli locali. Nonostante la censura del tempo, la stampa cominciò a svolgere un ruolo cruciale nella vita politica e sociale e in Olanda, dove la stampa incontrò meno impedimenti, si diffusero fogli chiamati "*corantos*" che fornivano notizie di carattere internazionale e riuscirono a circolare anche nei paesi più controllati. Fu in questo contesto che cominciò a emergere il concetto di "notizia", che non si limitava più a eventi straordinari o messaggi di natura commerciale, bensì comprendeva temi legati alla vita di tutti i giorni.

A metà del 17° secolo, accanto alle gazzette si diffusero anche le riviste, periodici specializzati in ambito culturale e letterario, di cui il primo esempio è

il "*Journal des savants*", che nacque a Parigi nel 1665, seguito poi dal "*Giornale dei letterati*" a Roma nel 1668. È sempre in questo periodo che uscì a Lipsia in Germania, in particolare nel 1660, il primo quotidiano della storia, intitolato "*Leipziger Zeitung*" con il sottotitolo "*Notizie fresche degli affari, della guerra e del mondo*" e qualche anno dopo nel 1702, uscì a Londra il quotidiano "*The Daily Courant*", pubblicato ininterrottamente per trentacinque anni, il quale divenne un modello per tutto il giornalismo europeo grazie alla sua rigorosa impostazione. Verso la fine del Settecento, il panorama della stampa si stava allargando e diffondendo in tutta Europa, ma solo nell'Ottocento, il giornalismo ha acquisito le caratteristiche moderne, beneficiando dello sviluppo tecnologico, del miglioramento delle vie di comunicazione e dell'evoluzione della società. Tra il 1830 e il 1850, nascono in Francia e in Inghilterra i principali giornali d'informazione, ma la vera rivoluzione avviene negli Stati Uniti con l'arrivo della penny press, che propone giornali ad un penny. Questo cambiamento è reso possibile dalla riduzione del costo della carta e dal supporto finanziario della pubblicità, permettendo una stampa in maggior numero di copie e una distribuzione a un prezzo più accessibile, rendendo così i quotidiani accessibili a tutti. Questo porta a un radicale cambiamento di contenuto, con notizie indirizzate a tutte le classi sociali, comprese cronache cittadine, nere e mondane e al consolidamento della figura del reporter, cacciatore di notizie. A New York, negli anni Trenta dell'Ottocento, nascono quotidiani come "*The New York Sun*", "*The New York Herald*", "*The New York Tribune*", fino al "*The New York Times*" nel 1851, che diventano modelli per il giornalismo mondiale. Anche in Europa, in questo periodo, vedono la luce i principali quotidiani ancora oggi in circolazione: in Francia "*Le Figaro*" (1826), in Inghilterra il "*Daily Telegraph*" (1855), in Germania la "*Frankfurter Zeitung*" (1856). In Italia, la stampa ha dovuto superare lunghe resistenze dovute alla censura, riuscendo a

raggiungere il livello degli altri paesi solo dopo l'unificazione nazionale. Verso la fine dell'Ottocento, nascono quotidiani di rilevanza come il *"Corriere della Sera"* (1876) a Milano, *"La Stampa"* nel 1895 a Torino, *"Il Messaggero"* (1878) a Roma, *"Il Mattino"* (1892) a Napoli, *"Il Resto del Carlino"* (1885) a Bologna, *"Il Gazzettino"* (1887) a Venezia. Verso la metà dell'Ottocento, vide la luce anche la prima generazione di agenzie internazionali di stampa, che si dedicarono alla diffusione delle notizie: la *"Havas"* in Francia, la *"Wolff"* in Germania e la *"Reuters"* in Inghilterra, istituzioni che hanno profondamente segnato la storia dell'informazione mondiale. In Italia, su spinta di Cavour, fu fondata l'agenzia Stefani, la quale successivamente, a partire dal 1945, assunse il nome di ANSA (Agenzia Nazionale Stampa Associata). Nel corso del Novecento, i giornali consolidano progressivamente il loro ruolo centrale nella società, esercitando una crescente influenza sulle decisioni politiche e sull'indirizzo dell'opinione pubblica (Giacheri Fossati L., 2005). Inoltre, in questo periodo, la pubblicità e la vendita di spazi promozionali divennero la principale fonte di reddito per i giornali, superando di gran lunga i proventi derivanti dalla distribuzione. Questa dipendenza dal settore pubblicitario ebbe un notevole impatto sulle dinamiche organizzative e produttive del giornalismo. La strategia di mercato predominante non risiedeva più nell'offerta di notizie ai lettori, ma nell'attrarre un pubblico per gli inserzionisti, dinamica che provocò pressioni sui giornalisti affinché raggiungessero un pubblico più vasto o specifico in base alle richieste degli inserzionisti (Siapera E., 2019).

Durante la Prima Guerra Mondiale, i giornali di tutti i paesi coinvolti vengono utilizzati come strumenti di propaganda patriottica e sono soggetti a censura preventiva, giustificata dal timore di un possibile sfruttamento delle notizie da parte del nemico. Questo controllo militare sull'informazione persiste anche nei conflitti successivi. Tuttavia, sono soprattutto i regimi totalitari del primo

Novecento, come il fascismo, il nazismo e i regimi comunisti, a sfruttare una stampa subordinata come mezzo per plasmare l'opinione pubblica e per imporre il consenso (Giacheri Fossati L., 2005).

2.2.1 La prima crisi

La prima rivoluzione che cambiò gli equilibri nel mondo del giornalismo verso la fine del Novecento, fu la diffusione della televisione, che divenne la fonte principale per la divulgazione delle notizie, e quindi diretta competitor dei giornali, con l'aggiunta anche di contenuti di intrattenimento. La televisione si distinse dagli altri mezzi di comunicazione non solo per la sua capacità di offrire notizie in tempo reale e coinvolgere il pubblico, ma soprattutto per la sua onnipresenza. Infatti, grazie alla diffusione sempre più ampia del segnale, alla sua accessibilità gratuita e alla capacità di catturare l'attenzione degli utenti in maniera decisamente incisiva, la televisione conquistò un vasto pubblico di telespettatori, consentendo una rapida diffusione delle informazioni a livello globale. L'avvento di questa trasformazione non solo ha portato nuove tecniche e professionisti nel campo del giornalismo, ma ha anche rivoluzionato il panorama dell'informazione, influenzando la centralità del giornale quotidiano e, di conseguenza, il modo stesso di praticare e diffondere il giornalismo. Un ulteriore elemento determinante per questa trasformazione è stata la diminuzione degli investimenti nella stampa a favore della televisione da parte delle agenzie pubblicitarie. Questo, unito ai crescenti costi di produzione non compensati dalla vendita dei giornali, ha scatenato la prima grande crisi nell'ambito del giornalismo cartaceo. In questo contesto, gli editori introdussero una trasformazione radicale all'interno del sistema giornalistico cartaceo, offrendo un servizio di qualità superiore, con una marcata differenziazione rispetto al prodotto televisivo, contrapponendosi alla sua immediatezza, sintomo di una trattazione superficiale delle notizie. Si procedette con la

modifica del layout del giornale, la sua struttura e la disposizione degli articoli. Il volume di pagine fu incrementato per consentire spazio ad analisi dettagliate e studi culturali più approfonditi e si estese l'area dedicata a nuovi temi, come lo sport, l'intrattenimento e la cronaca. Si trattò di una rivoluzione e di un considerevole investimento economico che consentì alla stampa cartacea di resistere alla crescente presenza della televisione (Brozzetti E. e Segatori R., 2015).

2.2.2 La seconda crisi e la nascita del giornalismo online

L'avvento del *World Wide Web* nel 1991 rappresentò un ulteriore elemento critico per il giornalismo scritto e le sue relazioni con il potere. Ideato da due informatici del CERN, il britannico Tim Berners-Lee e il belga Robert Cailliau, il *World Wide Web* nacque come un tentativo di creare un sistema che permettesse la pubblicazione e il recupero di informazioni online, favorendo la connessione tra diverse comunità scientifiche internazionali e modificando permanentemente non solo l'ambito tecnologico ma anche quello della comunicazione. La rete fornisce accesso a una vasta gamma di contenuti (testi, video, immagini) in qualsiasi momento e in qualsiasi parte del mondo, permettendo all'utente di gestire direttamente i suoi percorsi di lettura e ricerca grazie all'ipertestualità e alla multimedialità. Nel web, il protagonista principale è l'utente connesso che consuma e crea contenuti, dando una svolta rispetto ai media tradizionali in cui la comunicazione è unidirezionale, senza possibilità di partecipazione attiva da parte dell'utenza. I media digitali e le tecnologie di rete cambiano radicalmente questa dinamica.

L'era del giornalismo online ha origini relativamente recenti e affonda le sue radici nel contesto americano. È infatti nei primi anni '90 che fecero la loro comparsa le prime edizioni digitali di quotidiani come il *Chicago Tribune* e il *Mercury Center* (versione web del *San José Mercury News*, quotidiano della

Silicon Valley), riconosciuti ancora oggi come precursori del giornalismo online. A iniziare questa avventura nel mondo del web furono soprattutto le redazioni di piccoli giornali a carattere locale, che videro nel digitale l'opportunità di allargare il proprio pubblico e di estendere la propria portata, pur mantenendo costi di produzione e gestione contenuti. Il loro intento era di introdurre un formato originale basato su aggiornamenti e rubriche appositamente pensate per il web, affiancati da contenuti ripresi dall'edizione cartacea. Questo modello commerciale prevedeva una combinazione di contenuti gratuiti e sezioni a pagamento, inclusa la replica digitale dell'edizione su carta. Tuttavia, dal punto di vista economico, le strategie adottate non portarono a risultati soddisfacenti. L'alto costo degli abbonamenti, unito alla scarsa diffusione del digitale nel paese, generò un impatto quasi fallimentare (Brozzetti E. e Segatori R., 2015). A partire dal 1994, fecero la loro comparsa i quotidiani più rinomati. Nel giugno di quell'anno, il *New York Times* inaugurò il proprio sito web che offriva principalmente recensioni, forum di discussione e liste tematiche. Poco dopo, anche il *Washington Post* debuttò sulla rete, seguito dal *Wall Street Journal* nel '96.

Il giornalismo online in Italia fu avviato da un quotidiano locale, l'*Unione Sarda*, nell'estate del 1994. Il progetto venne portato avanti con grande speranza dal proprietario Nichi Grauso, avendo particolare attenzione anche per la struttura e la forma del sito, ma ebbe risultati diversi dal previsto e ben presto si ridimensionò notevolmente. I principali elementi contrari allo sviluppo del progetto furono la ridotta diffusione della rete allora presente in Italia e la natura regionale del quotidiano, che non permetteva una diffusione agevole a livello nazionale/globale. I primi grandi quotidiani sbarcarono nella primavera '95 e furono *La Stampa*, *Il Corriere della Sera* e *La Gazzetta dello Sport*, con siti semplicistici che riportavano notizie presenti già nelle versioni cartacee. Infatti,

in questa fase iniziale si seguiva la tecnica del *repurposing* che consisteva nel proporre sulla versione web, al minimo costo possibile, gli stessi contenuti del prodotto cartaceo, con l'obiettivo di ottenere guadagni aggiuntivi con i banner pubblicitari inseriti nelle pagine. Poco fortunato fu anche il tentativo dell'*Unità*, che ebbe inizio il 30 agosto 1995 e che seguiva le tecniche dei suoi predecessori. Sebbene ci fossero alcuni esempi di buon livello, oltre un anno dopo il debutto del primo progetto, il panorama del giornalismo online italiano risultava ancora poco soddisfacente, sia in termini di qualità che di completezza. I primi segni di ripresa furono nella primavera del 1996, in concomitanza con le elezioni politiche, quando *Repubblica* cominciò a delineare un potenziale ingresso nel mondo online, sperimentando con il sito *www.repubblica.interbusiness.it*, interamente focalizzato sull'evento elettorale. L'iniziativa, originata dalla visione innovativa di un redattore appassionato di informatica, Vittorio Zambardino, riscosse un successo straordinario per l'epoca, raggiungendo la notevole cifra di 350.000 visite in soli venti giorni. Dopo mesi di lavoro, il 14 gennaio 1997 fu ufficialmente lanciato *Repubblica.it*, l'estensione digitale del giornale cartaceo. Il sito, basato su alcune delle migliori pratiche dell'epoca, divenne immediatamente un punto di riferimento per il giornalismo online italiano e fu un veicolo per ampliare l'audience, il mercato e l'importanza culturale della testata. A differenza del precedente trend, *Repubblica.it* offriva contenuti aggiuntivi rispetto alla versione cartacea, permettendo agli utenti online di avere a disposizione un giornale telematico sempre aggiornato e con contenuti diversi rispetto all'edizione fisica. Nel corso del 1998, il numero di siti italiani che fornivano contenuti giornalistici ebbe una crescita di oltre il 200%, ma il vero picco del fenomeno si registrò nel 1999, considerato l'anno della svolta a causa della trasformazione radicale delle redazioni web, che subirono un forte ampliamento. Con l'avvento del 2000, il panorama del giornalismo

online italiano subì una completa trasformazione con l'entrata decisa dei principali gruppi editoriali che non avevano ancora sviluppato un sito competitivo su internet e si consolidò in modo definitivo (Bettini A., 2006).

	QUOTIDIANI	RIVISTE	WEBZINE
DICEMBRE 1997	44	226	128
DICEMBRE 1998	54	326	197
DICEMBRE 1999	62	501	412
DICEMBRE 2000	76	559	772
MARZO 2001	80	633	905

Fig. 6, Dati sul mercato giornalistico italiano dal 1997 al 2001. Fonte: Magrini, 2002

2.3 I giornali online come piattaforme digitali

I giornali rappresentano un tipico esempio di *multi-sided platform*, in quanto permettono l'interazione tra gli inserzionisti, i consumatori e i produttori di notizie, e generano effetti di rete indiretti, accoppiando la domanda che i lettori hanno per gli articoli di notizie, con la domanda che gli inserzionisti hanno per un prodotto completamente diverso, gli spazi pubblicitari ed un vasto pubblico. I prezzi per i consumatori dipendono dalla domanda degli inserzionisti per raggiungere quei consumatori e i prezzi per gli advertisers dipendono dal numero di utenti che riescono a raggiungere tramite la piattaforma. Generalmente, gli inserzionisti sovvenzionano il costo del consumo di articoli di notizie per i lettori quasi completamente (Nachbar T. B., 2021).

2.3.1 L'economia dell'attenzione

Gli utenti dei media sono coloro che alimentano la cosiddetta economia dell'attenzione e le loro decisioni su cosa leggere, guardare o condividere plasmano tale mercato. Tuttavia, molti altri fattori influenzano il comportamento del pubblico, e quindi dove l'attenzione viene raccolta e

monetizzata, come le strategie e politiche delle piattaforme e gli algoritmi utilizzati. In questo scenario quindi, l'attenzione gioca un ruolo essenziale per la sostenibilità del giornalismo ed è definita da Merja Myllylahti come *“una merce scarsa e fluida che porta valore monetario, la quale si basa sull'interazione individuale dell'utente che può essere raccolta, misurata e scambiata per ricavi su una piattaforma, come un sito di notizie”*. Tale interpretazione ne evidenzia tre dimensioni fondamentali approfondite di seguito: la limitatezza e la dinamicità, la misurazione e la quantificazione e, infine, la monetizzazione.

L'attenzione non è solo un concetto astratto, ma porta con sé un valore tangibile, che può essere convertito in ricavi. Il processo di mercificazione coinvolge le piattaforme nel trasformare attività, emozioni e idee, sia online che offline, in beni scambiabili e avviene tramite degli intermediari, che vendono l'attenzione umana a terze parti in cambio di ricavi, ovvero le piattaforme online, tra cui i giornali digitali, Facebook e Google. L'attenzione può essere considerata *“fluida”*, in quanto si muove costantemente tra diverse piattaforme e siti di notizie, influenzando le interazioni degli utenti e non può essere accumulata o conservata per un utilizzo futuro, ma deve essere costantemente rinnovata per poterne usufruire. Inoltre, il cervello umano ha limitata capacità di elaborazione, caratteristica che porta gli utenti a prendere costantemente delle *“decisioni di attenzione”*, scegliendo se considerare le notizie o meno e rendendo l'attenzione un bene scarso. Al fine di sfruttare l'attenzione, essa deve essere raccolta, analizzata e misurata, in quanto ogni forma di interazione dell'utente può essere trasformata in dati scambiabili per pubblicità o entrate da parte dei lettori attraverso abbonamenti digitali. Le aziende di informazione hanno dovuto adottare strumenti sempre più sofisticati per monitorare in tempo reale le interazioni dell'audience, in quanto le metriche risultanti sono diventate parte integrante delle operazioni delle redazioni, delle strategie

editoriali e delle valutazioni delle performance e la qualità dei dati ricevuti influisce sui ricavi, sulle decisioni su come e dove monetizzare i contenuti e su come investire in nuovi prodotti o personale. In questo contesto, i cookie svolgono un ruolo cruciale, in quanto consentono alle piattaforme di personalizzare l'esperienza dell'utente e di fornire contenuti mirati, ottimizzando così l'interazione e catturando l'attenzione in modo più efficace, e permettono di offrire agli inserzionisti dati sempre più accurati e validi. Secondo alcuni, questo tipo di "giornalismo misurabile" aggrega i comportamenti e gli atteggiamenti dei consumatori di notizie, incoraggiando i giornalisti a essere più reattivi nei confronti del loro pubblico. Tuttavia, è importante notare che strategie orientate ai "click" adottate in passato hanno eroso i loro valori editoriali, e perseguire l'attenzione comporta un rischio simile.

Prima dell'avvento di Internet, i giornali cartacei ottenevano considerevoli profitti sia grazie ai ricavi diretti provenienti dalle vendite sia grazie all'attenzione dei loro lettori, essendo l'unico canale per raggiungere il pubblico e detenendo quindi il monopolio nel mercato della pubblicità. La diffusione di Internet e delle piattaforme online ha destabilizzato il settore del giornalismo, cambiando radicalmente l'ordine consolidato intorno alle notizie attraverso una redistribuzione del potere nel mercato dell'advertising. Piattaforme come Facebook hanno assunto un ruolo dominante, con un vasto numero di utenti e la capacità di raccogliere informazioni dettagliate sugli stessi, rivendendole a inserzionisti e altri partner terzi per trarne profitto. Maggiore è la conoscenza che le piattaforme hanno dei propri utenti, maggiore è l'attrazione che esse hanno per gli inserzionisti, desiderosi di personalizzare e mirare con precisione le proprie campagne di marketing. Nel passaggio dall'ambiente stampato a quello digitale, le aziende mediatiche avevano la speranza di mantenere

l'attenzione dei lettori (Myllylahti M., 2020). Sebbene infatti nel modello tradizionale cartaceo i ricavi provenivano sia dalla vendita diretta di contenuti che dagli introiti pubblicitari, molti giornali inizialmente offrivano le edizioni online gratuitamente, monetizzandole solo in modo indiretto attraverso la pubblicità. Tuttavia, si sono trovate a dover competere strenuamente con le piattaforme emergenti e, nonostante gli sforzi profusi, non sono riuscite a generare entrate significative dalle loro audience digitali e a compensare le perdite di ricavi dalle fonti offline (Rußell R. et al, 2020). Gli studi accademici precedenti nel campo del giornalismo digitale e dei media più in generale, hanno indagato su quanto successo abbiano le aziende di informazione nel convertire l'attenzione degli utenti in ricavi e come interazioni individuali, come i "mi piace" e le condivisioni, possano essere trasformate in fonti di guadagno, senza però raggiungere il valore monetario preciso dell'attenzione ottenuta sulle piattaforme. La condivisione di contenuti e post di notizie sui social media può innalzare il traffico verso i siti web dei giornali online e aumentare il numero di visualizzazioni, il che risulta allettante per gli inserzionisti. Tuttavia, sebbene le aziende di informazione abbiano ottenuto un notevole coinvolgimento e traffico tramite le piattaforme di social media, i benefici monetari non sono stati altrettanto significativi, rendendo dubbia la strategia di distribuzione di notizie attraverso tali canali. Se le aziende di informazione aderiscono ai servizi di abbonamento offerti dai social media e agli accordi di condivisione dei ricavi, la loro dipendenza da questi ultimi crescerà in misura ancora maggiore (Myllylahti M., 2020). Nei modelli digitali è stato quindi introdotto il concetto di *paywall*, un sistema mirato a limitare o controllare l'accesso gratuito ai contenuti, richiedendo agli utenti di effettuare un pagamento o registrarsi per ottenere accesso completo alla piattaforma. L'obiettivo dei *paywall* è quello di generare entrate dirette dai consumatori, che

contribuiscono a sostenere l'operatività e la produzione. A causa della costante concorrenza delle alternative di contenuti gratuiti su Internet, i consumatori mostrano una generale riluttanza a pagare per i contenuti online rispetto alle versioni fisiche, creando incertezza nella corretta tariffazione per l'accesso ai contenuti digitali. Per il pricing delle versioni online, è importante considerare che un prezzo troppo alto potrebbe ridurre la fruizione, agendo sia sul versante dei ricavi diretti che sul versante dei ricavi pubblicitari. Pertanto, è strettamente necessario configurare con attenzione i paywall al fine di ottimizzare la monetizzazione dei contenuti (Rußell R. et al, 2020). Il pioniere di questa strategia fu Rupert Murdoch, probabilmente il titano dei media più influente del suo tempo, che deteneva, tra gli altri, *The Sun*, *The Times* e *The Wall Street Journal*. Nel 2008, annunciò l'intenzione di avviare una politica di abbonamento per i siti dei suoi giornali, sostenendo *"La stampa sta attraversando un'epocale dibattito sull'opportunità di dare accesso ai propri siti gratis o a pagamento. Dalla nostra esperienza al Wall Street Journal è ovvio che è possibile far pagare. L'era attuale di internet"*, intesa come un'era in cui l'informazione giornalistica viene data per lo più gratuitamente, *"sarà presto finita"* (Franceschini E., 2009). In Italia, è stato necessario attendere sette anni prima che questo approccio fosse adottato. Il 25 gennaio 2016, il Corriere della Sera ha comunicato l'intenzione di rendere i propri articoli disponibili esclusivamente a pagamento. Come dichiarato in una nota ufficiale, a partire dal 27 gennaio *"i lettori avranno accesso gratuito, su pc, tablet e smartphones, a 20 articoli al mese. Dopo questa soglia verrà chiesto di contribuire con un abbonamento. L'offerta base – 9,99 euro al mese, solo 0,99 euro per il primo mese – dà accesso a tutti i contenuti del sito, da desktop e da mobile"* (Ferrazza F., 2016).

2.4 Il mercato attuale italiano

Nel contesto del giornalismo italiano, l'analisi dei principali attori editoriali attraverso indicatori chiave come il numero di visitatori mensili, pagine visualizzate e dati finanziari fornisce la prospettiva sullo stato e l'evoluzione del settore. Questo approccio permette di valutare l'impatto delle testate giornalistiche più influenti, offrendo una panoramica dettagliata sulle loro performance online e sulle dinamiche economiche. Di seguito, saranno esplorate alcune delle metriche e dei dati più rilevanti per comprendere il panorama del giornalismo italiano e le sfide che il settore affronta nell'era digitale.

2.4.1 Panoramica generale

Partendo da una panoramica generale, secondo i nuovi dati di *Audipress* 2023 che coprono il periodo da settembre 2022 a luglio 2023, in Italia si contano 31,9 milioni di individui che hanno avuto contatti con un titolo stampato su carta e/o in formato digitale, raggiungendo il 61,1% della popolazione dai 14 anni in su. Le copie in formato digitale contano 7,1 milioni di lettori totali di cui 2 milioni e 462 mila di quotidiani, confermando la tendenza positiva e costante che ha preso avvio durante il periodo pandemico. Questo fenomeno infatti è esploso a partire dai dati raccolti durante il 2020, quando si registravano 3,2 milioni di lettori di copie online, rispetto all'edizione precedente (2019) che ne contava solo 1,5 milioni, e ha raggiunto ad oggi un segmento stabile e fedele di lettori. Passando ai dettagli sui formati e le modalità di lettura delle testate prese in esame, emerge che i Quotidiani contano una media di 11,7 milioni di lettori su carta e/o in formato digitale nel giorno medio (equivalente al 22,5% della popolazione dai 14 anni in su), generando una media di 17 milioni e 140 mila letture al giorno.

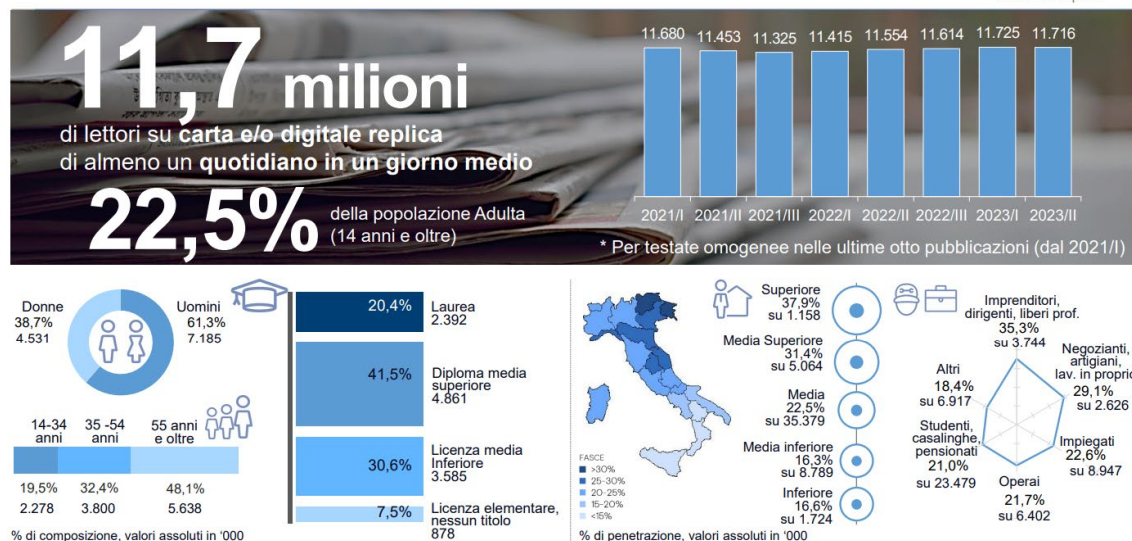


Fig. 6, I lettori dei quotidiani in Italia. Fonte: Audicom, 2023

2.4.2 Panoramica per singola testata

Proseguendo con l'analisi, si evidenziano i dati di audience relativi ai principali siti online del mercato delle principali testate italiane. I grafici che seguono sono stati generati a partire dalle statistiche fornite dal sito "Audiweb" relative all'audience dei siti di informazione online nei primi sette mesi del 2023 (da gennaio a luglio), rielaborati calcolando la media mensile di utenti unici e pagine visitate mensilmente per ciascuna pubblicazione. Attraverso tali grafici, sarà possibile mettere in luce l'efficacia e l'attrattività delle diverse pubblicazioni nel panorama digitale nazionale, consentendo di cogliere tendenze e variazioni significative tra i diversi attori. Di seguito, si propone la legenda delle due variabili prese in esame proposta da *Audiweb* per una maggiore comprensione.

- 1) **UTENTI UNICI:** numero di singole persone, non replicate che si sono collegate in un determinato arco temporale ad un sito e/o elementi di esso effettuando una o più visite. Il sistema Audiweb consente di riconoscere un utente unico anche quando:

- cancella o blocca i cookies;
- accede da diversi device;
- naviga con più di un browser;
- accede da un altro computer.

2) PAGINE VISTE: numero totale di pagine viste e di App views.

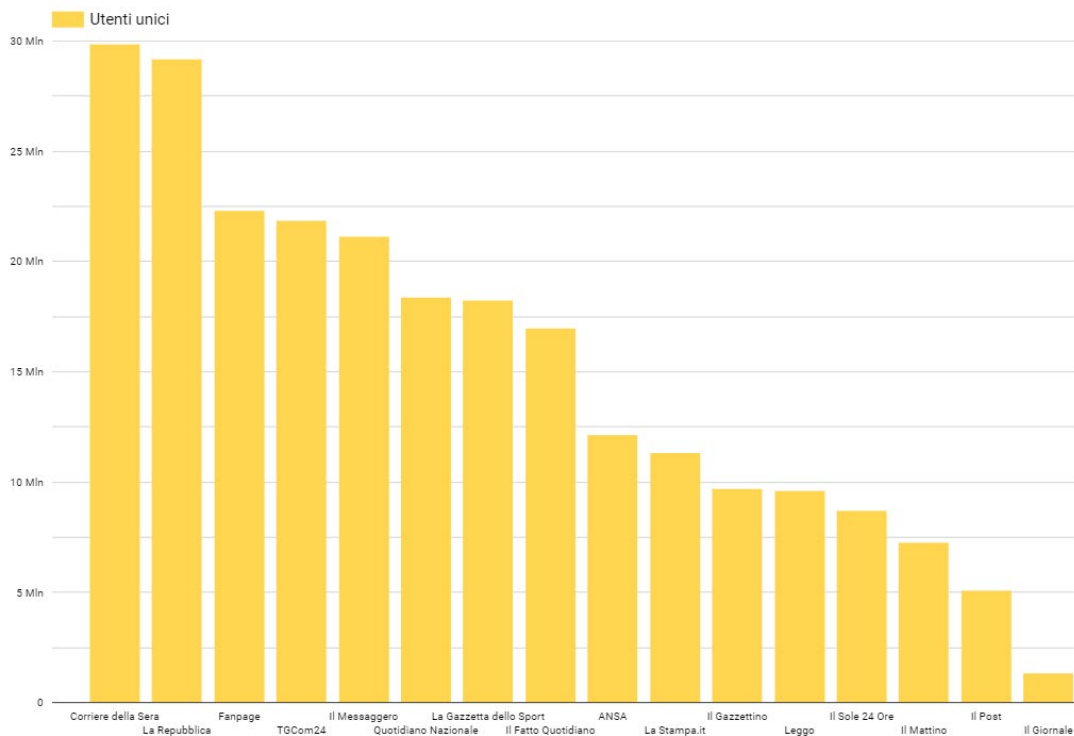


Fig. 7, Numero medio mensile di utenti unici. Fonte: Elaborazione personale

La Repubblica e il *Corriere della Sera* emergono come leader incontrastati, registrando un numero di utenti che si aggira intorno ai 30 milioni, seguiti da *Fanpage*, *TGCom24* e *Il Messaggero* che superano la soglia dei 20 milioni. Questo suggerisce una presenza online estremamente forte e un ampio coinvolgimento della loro audience. Inaspettatamente, alla fine della classifica si posizionano testate importanti come *Il Sole 24 Ore* e *Il Post* con un numero di utenti unici significativamente inferiore che non tocca i 10 milioni di utenti, indicando una presenza online più limitata.

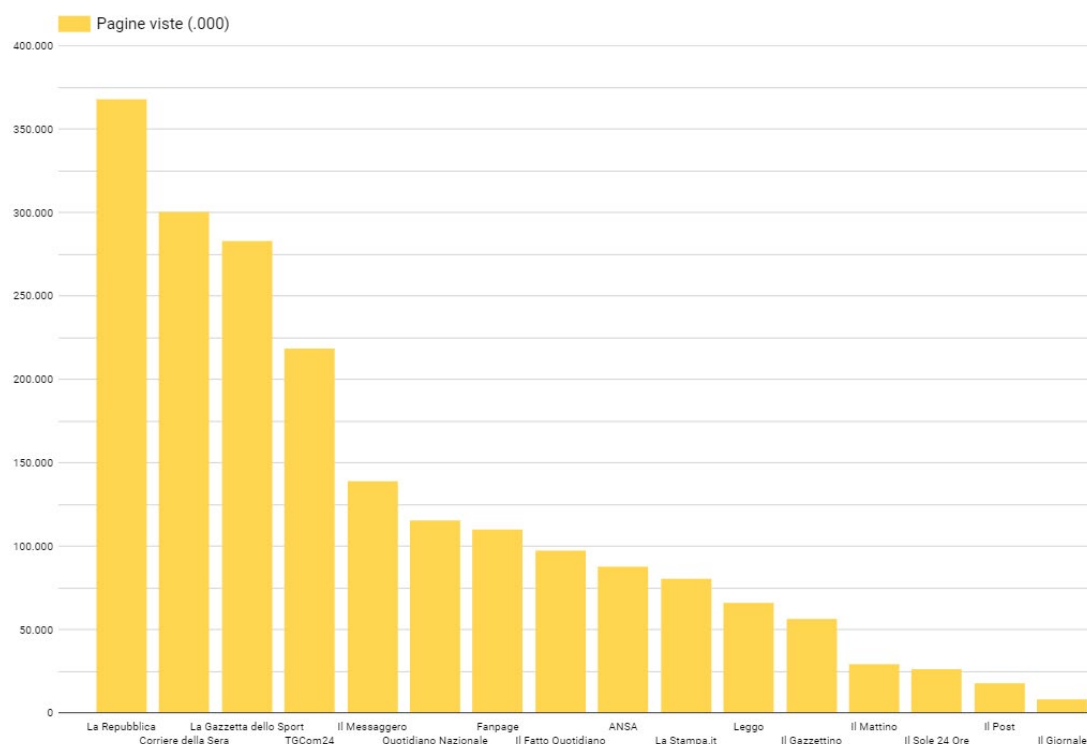


Fig. 8, Numero medio mensile di pagine viste. Fonte: Elaborazione personale

Anche per numero di pagine visitate, *La Repubblica* e il *Corriere della Sera* continuano ad essere nelle prime posizioni, confermano l'ampia portata dei loro contenuti e l'alto coinvolgimento degli utenti che navigano nei loro siti. Si riconfermano agli ultimi posti *Il Sole 24 Ore* e *Il Post*, mostrando un numero significativamente inferiore di pagine viste.

In generale, le due classifiche risultano essere abbastanza coerenti tra loro mostrando un quadro delineato del mercato attuale.

2.4.3 Panoramica finanziaria del mercato dell'editoria

L'analisi dei principali indicatori di bilancio nel mercato dell'editoria e delle testate giornalistiche italiane riveste un ruolo cruciale nell'offrire una rappresentazione accurata e dettagliata dello stato finanziario e operativo di questo settore. A tal proposito, si riporta il “*Rapporto AGCOM Focus Bilanci (2017-2021)*” sul mercato dell'editoria, che ha incluso le seguenti imprese: *Altroconsumo, Arnoldo Mondadori, Athesis, Avvenire, Bresi, Cairo Communications,*

Caltagirone Editore, Class, De Agostini Editore, Edizioni Condè Nast, Edizioni Il Fatto, Gruppo Editoriale l'Espresso, Hachette, Hearst magazine, Monrif, Panini, Periodici S. Paolo, RCS Mediagroup, Sesaab, Sole 24 Ore, Wolters Kluwert.

La prima voce analizzata è quella dei ricavi. Nel periodo preso in considerazione, si osserva un calo aggregato del 10,0% nei ricavi delle principali imprese attive nel settore dell'editoria quotidiana e periodica. Questi scendono da 4,24 miliardi di euro nel 2017 a 3,82 miliardi di euro nel 2021. I ricavi domestici subiscono una flessione dell'8,4%, passando da 3,76 a 3,44 miliardi di euro. Tuttavia, va notato che questo valore nel 2021 ha registrato una crescita del 7% rispetto al 2020.

Guardando invece alle diverse fonti di ricavo nel mercato italiano, nel 2021 sono evidenziate tre principali voci:

- 1) Proventi Editoriali: rappresentano il 39,1% del totale dei ricavi. Si riferiscono ai guadagni derivanti dalla vendita di prodotti editoriali, come giornali, riviste, libri e altre pubblicazioni simili. Questo può includere sia le vendite in edicola che le sottoscrizioni, sia in formato cartaceo che digitale;
- 2) Ricavi Pubblicitari: costituiscono il 30,5% del totale dei ricavi. Si tratta dei guadagni ottenuti dalla vendita di spazi pubblicitari all'interno delle pubblicazioni su giornali, riviste, siti web e altri mezzi di comunicazione;
- 3) Editoria Libreria: rappresenta il 13,7% dei ricavi totali. Si riferisce ai guadagni derivanti dalla vendita di libri, fisici o digitali.

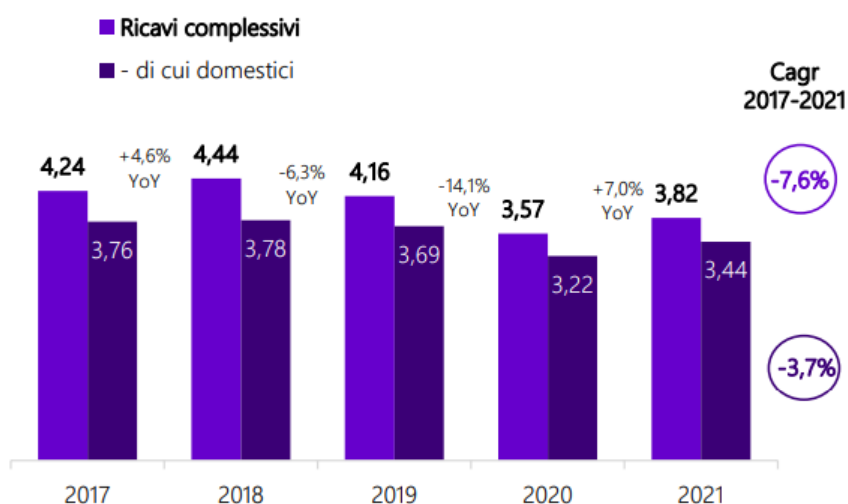


Fig. 9, Andamento ricavi 2017-2021. Fonte: AGCOM



Fig. 10, Composizione ricavi 2021. Fonte: AGCOM

Estendendo l'analisi al periodo compreso tra il 2012 e il 2021 per quanto concerne le principali fonti di reddito, ovvero i proventi editoriali e i ricavi pubblicitari nel mercato domestico, emergono valutazioni aggregate che evidenziano una significativa contrazione. Nel dettaglio, nel corso di questo decennio, i proventi editoriali sono scesi del 38,5%, passando da 1,86 miliardi di euro a 1,14 miliardi di euro, mentre i guadagni derivanti dalla vendita di spazi pubblicitari hanno subito una contrazione del 45%, passando da 1,78 miliardi di euro a 0,98 miliardi di euro. Se concentriamo l'attenzione sui due sottoperiodi distinti, 2012-2016 e 2017-2021, emerge un trend simile: sia i proventi editoriali

che i ricavi pubblicitari mostrano una riduzione cumulativa. Nel dettaglio, i ricavi da attività editoriale registrano una diminuzione del 14,8%, passando da 7,7 miliardi di euro a 6,6 miliardi di euro, mentre gli introiti pubblicitari subiscono un decremento del 31,0%, scendendo da 7,3 miliardi di euro a 5,1 miliardi di euro.

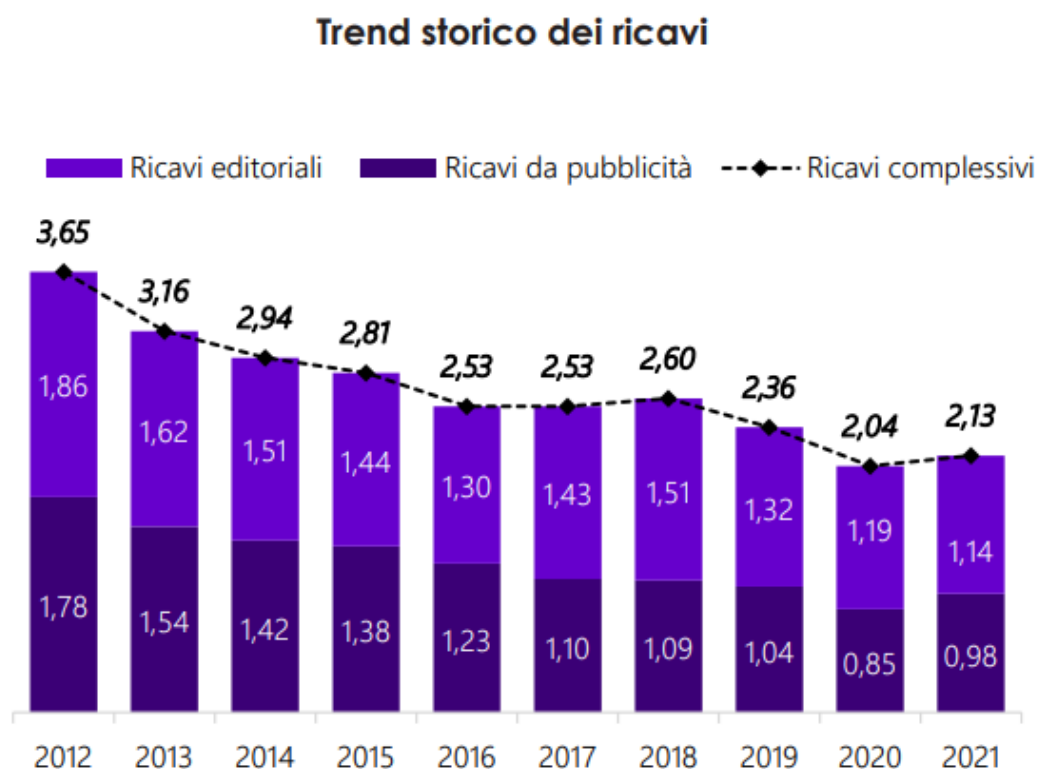


Fig. 11, Trend ricavi editoriali e pubblicitari 2012-2021. Fonte: AGCOM

Nonostante un trend di introiti in diminuzione, le strategie di ristrutturazione attuate dalle aziende, che hanno comportato una riduzione media del 43,2% nei costi operativi e quelli relativi al personale, hanno avuto un impatto positivo sui margini operativi. Complessivamente, il margine lordo nel periodo preso in esame ammonta a 3,60 miliardi di euro. Tuttavia, se analizziamo i sottoperiodi 2012-2016 e 2017-2021, i dati aggregati risultano rispettivamente di 1,5 miliardi di euro e 2,1 miliardi di euro, evidenziando un miglioramento del 35% nel

secondo quinquennio. Relativamente ai ricavi, l'indice passa dall'6,2% all'11,1% (con un 8,3% nell'intero periodo). Il margine netto (Ebit) segue una tendenza simile: sebbene nel complesso si registri una perdita di circa 0,70 miliardi di euro nell'intero periodo, nei due periodi considerati si osserva un notevole progresso. Il margine netto passa da -1,12 miliardi di euro a +0,42 miliardi di euro (dal -4,5% al +2,3% dei ricavi).

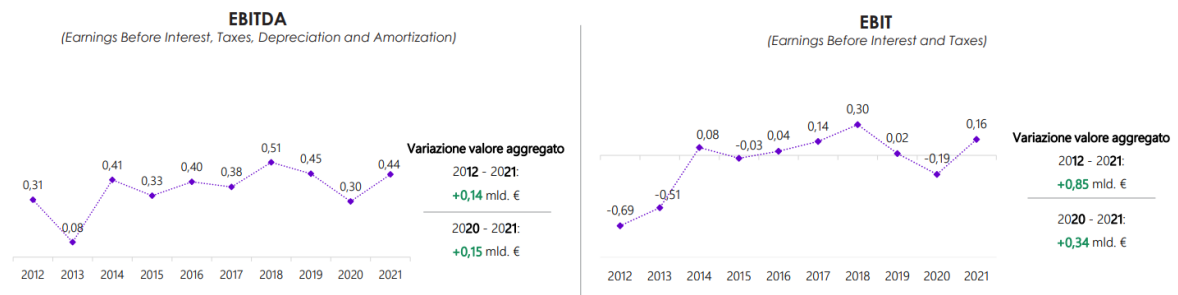


Fig. 12, Trend EBITDA e EBIT 2012-2021. Fonte: AGCOM

In ultima istanza, considerando il periodo 2016-2020, il flusso finanziario generato dalle attività operative, rapportato in percentuale ai ricavi, ha mediamente raggiunto il 9,8%. Si è registrata una crescita significativa tra il 2016 e il 2018, passando dal 5,4% all'11,1%. Tale valore ha subito una leggera diminuzione nel 2019 (9,8%), per poi risalire all'11,6% nel 2020.

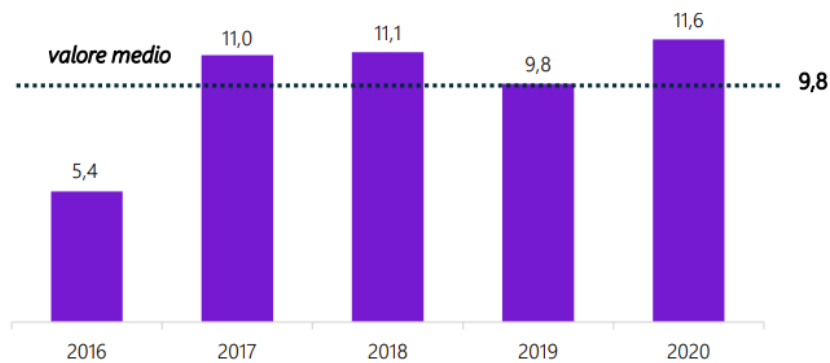


Fig. 13, Rapporto Cash Flow su Ricavi 2016-2020. Fonte: AGCOM

CAPITOLO 3 – ANALISI SUGLI EFFETTI DEL COOKIE WALL

3.1 L'introduzione del cookie wall

A partire dagli ultimi mesi del 2022, diverse testate giornalistiche online e siti di informazione italiani, hanno implementato strategie dette di *pay or consent* che vincolano l'accesso ai contenuti, offrendo agli utenti due opzioni: la sottoscrizione di un abbonamento (il cosiddetto paywall) o il rilascio del consenso all'installazione di cookie e altri strumenti di tracciamento dei dati personali (il cosiddetto cookie wall). La semplice visualizzazione dell'home page del sito (senza la possibilità di leggere il testo degli articoli) è così vincolata all'accettazione dei cosiddetti cookie di profilazione nel caso in cui non si possieda un abbonamento.



Fig. 14, Banner informativa cookie, La Repubblica. Fonte: www.larepubblica.it

Come già illustrato in precedenza, i cookie sono stringhe di testo che i siti web visitati dagli utenti inseriscono e memorizzano all'interno del loro dispositivo terminale. Questi sono utilizzati per una varietà di scopi, tra cui l'autenticazione informatica, il monitoraggio delle sessioni, la conservazione di informazioni sulle configurazioni specifiche degli utenti che accedono al server, la memorizzazione delle preferenze e la facilitazione della fruizione dei contenuti online. In base alla loro funzione, i cookie possono essere suddivisi in due categorie principali:

- 1) i cookie "tecnici", finalizzati a facilitare la navigazione o a fornire un servizio richiesto direttamente dall'utente. Di norma, vengono installati dal titolare del sito web senza richiedere un consenso esplicito da parte dell'utente. Senza l'utilizzo di tali cookie, alcune operazioni potrebbero risultare impossibili o più complesse e meno sicure. Ad esempio, nel caso delle attività di home banking (come la visualizzazione dell'estratto conto, bonifici o il pagamento di bollette), i cookie svolgono un ruolo essenziale consentendo l'identificazione e il mantenimento dell'utente all'interno della sessione;
- 2) i cookie "di profilazione", utilizzati per analizzare il comportamento dell'utente e creare un profilo dettagliato, permettendo di inviare pubblicità mirate, valutare l'efficacia delle campagne pubblicitarie e implementare strategie commerciali in linea con le preferenze dell'utente. Per l'uso di cookie di profilazione, il titolare del sito deve ottenere un consenso esplicito dall'utente. Quest'ultimo deve fornirlo in modo libero, specifico, informato ed inequivocabile, in accordo con gli articoli del GDPR.

Prima della policy, per accedere ai siti e fruire dei medesimi contenuti era possibile accettare solo i cookie tecnici, mentre con la sua introduzione, l'utente,

se non abbonato, deve necessariamente prestare il consenso anche per quelli di profilazione. Rimangono in ogni caso riservati ai soli abbonati determinati contenuti premium.

Tale policy mira a mitigare una evidente diminuzione dei ricavi derivanti dalla vendita di pubblicità mirata, per la quale gli inserzionisti sono disposti a investire molto di più rispetto a quella rivolta ad un pubblico generico. Questa crisi è stata causata da una serie di iniziative atte ad aumentare la consapevolezza degli utenti sull'attività di profilazione e sui loro dati in generale. Tra queste, emerge l'introduzione da parte del Garante dell'obbligo di utilizzare cookie banner che permettono all'utente, in modo trasparente, di accettare solo i cookie strettamente necessari.

Il tema in questione suscita divisione di opinioni. Da un lato, ci sono coloro che sostengono la monetizzazione dei dati personali, percependolo come un modo per ottenere un servizio migliore attraverso la profilazione, traducendosi spesso in pubblicità mirata su argomenti di interesse e comprendendo l'esigenza dei siti online di sostenere il loro lavoro anche attraverso questo canale. Dall'altro lato, ci sono coloro che sottolineano la mancanza di libertà e trasparenza nel consenso fornito. L'utente è libero di rifiutare in qualsiasi momento, ma in tal caso gli viene richiesto di sottoscrivere uno degli abbonamenti offerti. Inoltre, l'utente potrebbe non avere una conoscenza dettagliata della natura dei cookie o di tecnologie simili che il titolare chiede di installare tramite il banner (Iadecola L., 2022).

3.1.1 La posizione delle Autorità Garanti Europee

Le Autorità Garanti per la protezione dei dati personali concordano nel ritenere che il consenso fornito dall'utente in presenza unicamente del cookie wall non possa essere considerato validamente espresso, poiché rappresenta una scelta

obbligata. Infatti, in questa circostanza, per accedere alla pagina web, l'utente non ha altra opzione se non quella di accettare tutti i cookie. In particolare, l'*European Data Protection Board* ha confermato nelle sue Linee guida sul consenso che l'uso del cookie wall non costituisce una forma valida di prestazione del consenso, così come il Garante italiano ha affermato, nella guida sui cookie e altri strumenti di tracciamento, l'illegalità di tale strumento, a meno che non venga offerta all'interessato un'opzione alternativa per accedere al contenuto senza dover prestare il proprio consenso al trattamento dei dati. Sembra essere questo il caso della politica in esame, in cui il consenso potrebbe ritenersi liberamente prestato trattandosi di una scelta alternativa e non necessaria. In ogni caso la questione sulla validità del consenso espresso con la presenza di paywall e cookie wall è da ritenersi più controversa e le Autorità Garanti europee non hanno ancora preso una posizione ufficiale, mentre diversi garanti nazionali hanno esposto le proprie considerazioni. In particolare le autorità danese, austriaca e francese hanno sottolineato che il consenso per l'adozione di tali sistemi è considerato valido a condizione che l'alternativa offerta non limiti in modo significativo la libera espressione dello stesso. Anche l'Autorità spagnola per la protezione dei dati ha manifestato indirettamente la propria posizione, allineandosi agli altri Garanti europei, suggerendo che il cookie wall può essere implementato purché agli utenti sia chiaramente comunicato che hanno due opzioni per accedere al servizio: accettare l'uso dei cookie oppure optare per un'altra alternativa, non necessariamente gratuita, e che non richiede il consenso per l'uso dei cookie. Il Garante tedesco si è recentemente occupato di un caso che ha coinvolto una delle principali piattaforme di divulgazione di contenuti online del Paese. Inizialmente, l'Autorità ha dichiarato l'illegittimità dei sistemi di cookie wall/paywall presenti sulla piattaforma. Dato che l'editore in questione ha attività sia in Austria che in

Germania, il ricorso a tale decisione è stato presentato sia all'Autorità tedesca che a quella austriaca. Alla fine, in linea con quanto già stabilito dall'Autorità austriaca, anche il Garante tedesco ha considerato valido il consenso espresso in questa circostanza, a condizione che l'utente sia in grado di esprimerlo in modo attivo e trasparente. In Italia, il Garante non ha ancora espresso un'opinione su questa tematica. In accordo con le direttive della normativa sulla privacy e in sintonia con le decisioni adottate dagli omologhi europei, è probabile che anche in Italia l'uso del cookie wall in abbinamento con il paywall sia considerato ammissibile. Attualmente, l'Autorità è al lavoro per valutare attentamente la questione, come indicato nelle comunicazioni diffuse tra ottobre e novembre 2022. Tale valutazione mira a determinare se siano necessari interventi in linea con l'attuale contesto normativo (Fiore C., 2023).

3.2 Introduzione all'analisi e metodologia di costruzione del dataset

Il principale obiettivo del presente studio è quello di indagare gli effetti della policy precedentemente introdotta che prevede l'introduzione del cookie wall da parte di diversi siti online. In particolare, gli aspetti presi in considerazione in questa sede sono l'efficacia dell'advertising online e diverse variabili di engagement. A tal proposito, è stato elaborato ed utilizzato un dataset, illustrato nel paragrafo corrente.

La fase preliminare alla costruzione del dataset ha incluso l'identificazione dei maggiori siti di informazione italiani e la verifica empirica della loro adesione alla policy di interesse. Successivamente, per tutte le piattaforme che hanno adottato la strategia *pay or consent*, è stata ricercata la data di implementazione della stessa attraverso due metodi distinti: il contatto diretto attraverso mail con i siti interessati dove possibile o, in alternativa, l'impiego del software *Wayback*

Machine, un archivio digitale che consente di accedere a versioni precedenti di siti web, consentendo così di recuperare informazioni storiche sulle politiche di utilizzo dei cookie. In alcuni casi, non è stato possibile ottenere l'informazione con nessuno dei due metodi. La prima tabella che segue contiene tutti i siti web in cui è stata riscontrata la presenza della policy e le relative informazioni, mentre la seconda tutti i siti che non l'hanno adottata.

GRUPPO	DOMINIO	DATA IMPLEMENTAZIONE POLICY	FONTE
GEDI	lastampa.it	17/10/22	Email dal gruppo
	repubblica.it	17/10/22	Email dal gruppo
	ilsecoloxix.it	17/10/22	Email dal gruppo
	laprovinciapavese.geolocal.it	17/10/22	Email dal gruppo
	lasentinella.geolocal.it	17/10/22	Email dal gruppo
	huffingtonpost.it	17/10/22	Email dal gruppo
RCS	corriere.it	17/10/22	Email dal gruppo
	gazzetta.it	17/10/22	Email dal gruppo
	bergamo.corriere.it	17/10/22	Email dal gruppo
	corrieredibologna.corriere.it	17/10/22	Email dal gruppo
	brescia.corriere.it	17/10/22	Email dal gruppo
	corrierefiorentino.corriere.it	17/10/22	Email dal gruppo
	milano.corriere.it	17/10/22	Email dal gruppo
	roma.corriere.it	17/10/22	Email dal gruppo
	napoli.corriere.it	17/10/22	Email dal gruppo
	torino.corriere.it	17/10/22	Email dal gruppo
	corrieredelveneto.corriere.it	17/10/22	Email dal gruppo
	oggi.it	17/10/22	Email dal gruppo
	amica.it	17/10/22	Email dal gruppo
	iodonna.it	17/10/22	Email dal gruppo
	ilfattoquotidiano.it	05/10/22	Email dal gruppo
	ilsole24ore.com	10/11/22	Wayback Machine
	ansa.it	07/03/23	Wayback Machine
	tuttosport.com	01/03/23	Email dal sito
	corrieredellosport.it		
	qds.it		
GRUPPO SAE	iltirreno.it		
	lanuovaferrara.it		
	gazzettadimodena.it		
	gazzettadireggio.it		
	lanuovasardegna.it		

Fig. 15, Siti che hanno implementato la policy. Fonte: Elaborazione personale

GRUPPO	DOMINIO	GRUPPO	DOMINIO
ATHESIA	dolomiten.it	GEDI	nationalgeographic.it
	altoadige.it		limesonline.com
	giornaletrentino.it		lescienze.it
	ladige.it		quotidiano.net
Citynews	agrigentonotizie.it	MONRIF	ilrestodelcarlino.it
	anconatoday.it		ilgiorno.it
	arezzonotizie.it		lanazione.it
	avellinotoday.it		iltelegrafolivorno.it
	baritoday.it	SESAAB	ecodibergamo.it
	bolognatoday.it		laprovinciadilecco.it
	bresciatoday.it		laprovinciadisonorio.it
	brindisireport.it		laprovinciadico.it
	casertanews.it	ATHESIS	ilcittadinomb.it
	cataniatoday.it		larena.it
	cesenatoday.it		ilgiornaledivivenza.it
	chietitoday.it		bresciaoggi.it
	ferraratoday.it	GRUPPO CORRIERE	corrieredellumbria.corr.it
	firenzetoday.it		corrierediarezzo.corr.it
	foggiatoday.it		corrieredisiena.corr.it
	forlìtoday.it		gazzettadelsud.it
	frosinonetoday.it	Società Editrice Sud (SES)	gds.it
	genovatoday.it		ilpost.it
	ilpescara.it		ilgiornale.it
	ilpiacenza.it		lagazzettadelmezzogiorno.it
	latinatoday.it		espresso.repubblica.it
	lecceprima.it		iltempo.it
	leccotoday.it		ilmanifesto.it
	livornotoday.it		ildubbio.news
	messinatoday.it		fanpage.it
	milanotoday.it		avvenire.it
	modenatoday.it		italiaoggi.it
	monzatoday.it		liberoquotidiano.it
	napolitoday.it		laverita.info
	novaratoday.it		laprovinciacr.it
	padovaoggi.it		giornaledibrescia.it
	palermotoday.it		liberta.it
	parmatoday.it		gazzettadiparma.it
	perugiatoday.it		ilcentro.it
	pisatoday.it		quotidianodelsud.it
	pordenonetoday.it		lasicilia.it
	quicomo.it		unionesarda.it
	ravennatoday.it		ladiscussione.com
	reggiotoday.it		secoloditalia.it
	riminitoday.it		loccidentale.it
	romatoday.it		ildenaro.it
	salernotoday.it		lidentita.it
	sondriotoday.it		vocedimantova.it
	ternitoday.it		tgcom24.mediaset.it
	torinotoday.it		focus.it
	trentotoday.it		wired.it
	trevisotoday.it		donnamoderna.com
	triestepima.it		vocerepubblicana.it
	udinetoday.it		affaritaliani.it
	veneziatoday.it		agoravox.it
	veronasera.it		blitzquotidiano.it
	vicenzatoday.it		bsnews.it
	viterbotoday.it		linkiesta.it
	today.it		ilprimatonazionale.it

Fig. 16, Siti che non hanno implementato la policy. Fonte: Elaborazione personale

Si riporta infine il caso del gruppo CEDS (Caltagirone) che ha adottato una particolare variante della policy: nel momento in cui non si è abbonati e si non presta il consenso ai cookie, appare al fondo dei siti del gruppo il banner in figura 17, che rende meno agevole la navigazione del sito ma non blocca i contenuti. Il banner rimane finché l'utente non accetta i cookie o non si abbona al servizio.

Hai scelto di non accettare i cookie

La pubblicità personalizzata è un modo per supportare il lavoro della nostra redazione, che si impegna a fornirti ogni giorno informazioni di qualità. Accettando i cookie, ci aiuterai a fornire una informazione aggiornata ed autorevole.

In ogni momento puoi modificare le tue scelte tramite il link "[preferenze cookie](#)".

ACCETTA COOKIE

oppure

ABBONATI a partire da 1€

Fig. 17, Banner informativa cookie, Il Messaggero. Fonte: www.ilmessaggero.it

GRUPPO	DOMINIO	DATA IMPLEMENTAZIONE POLICY	FONTE
CEDS/ Caltagirone Editore	ilmessaggero.it	01/09/22	Email dal gruppo
	ilmattino.it	01/09/22	Email dal gruppo
	ilgazzettino.it	01/09/22	Email dal gruppo
	corriereadriatico.it	01/09/22	Email dal gruppo
	quotidianodipuglia.it	01/09/22	Email dal gruppo
	leggo.it	01/09/22	Email dal gruppo

Fig. 18, Siti del gruppo CEDS. Fonte: Elaborazione personale

Per la fase di costruzione, attraverso un data provider, sono stati ottenuti due set di dati completi riguardanti i siti di informazione più frequentati in Italia. In particolare, il primo set di dati è riferito all'efficacia dell'advertising ed è composto dalle seguenti variabili:

- 1) DOMINIO: sito a cui si riferiscono le informazioni;
- 2) DATA: data a cui si riferiscono le rilevazioni. In questo caso l'informazione temporale ha dettaglio mensile. Per ogni dominio si hanno rilevazioni che vanno da settembre 2020 a settembre 2023 per un totale di 37 osservazioni per sito;

- 3) OUTGOING AD TRAFFIC: click sugli annunci pubblicitari presenti sul sito web, che reindirizzano l'utente ad una pagina di destinazione esterna.

Il secondo dataset contiene informazioni sull'engagement dei siti di interesse ed è composto dalle seguenti variabili:

- 1) DOMINIO: sito a cui si riferiscono le informazioni;
- 2) DATA: data a cui si riferiscono le rilevazioni. In questo caso l'informazione temporale ha dettaglio giornaliero. Per ogni dominio si hanno rilevazioni che vanno dal 01/08/2020 al 03/10/2023 per un totale di 1158 osservazioni per sito;
- 3) VISITS: numero totale di visite al sito. Ogni visita può comprendere più pagine visualizzate da un singolo utente;
- 4) UNIQUE VISITATORS: numero totale di visitatori unici che hanno visitato il sito. Gli utenti unici vengono conteggiati una sola volta, indipendentemente da quante visite effettuano;
- 5) AVG. VISIT DURATION (o DURATION): durata media di una visita sul sito web;
- 6) PAGES/VISIT: numero medio di pagine visualizzate da un visitatore durante una visita;
- 7) BOUNCE RATE: percentuale di visite in cui l'utente ha visualizzato solo una pagina del sito web e poi ha lasciato senza interagire ulteriormente;
- 8) TOTAL PAGE VIEWS: numero totale di pagine visualizzate su tutto il sito web durante un determinato periodo di tempo.

Su entrambi i dataset sono state svolte operazioni di *pre-processing* attraverso Excel e l'utilizzo di codici VBA in quanto nei dati erano presenti *missing values* (N/A) e/o valori non ammissibili per l'analisi, tra cui valori uguali a "<5000"

senza informazione puntuale o pari a 0,00. Negli algoritmi usati sono state fissate delle soglie di accettazione riguardanti la presenza di tali valori per non eliminare un'eccessiva quantità di siti. In particolare, per il dataset sull'engagement con dettaglio giornaliero, la soglia di accettazione riguarda la somma tra valori "N\A" e valori "<5000" presenti, la quale è stata fissata a circa il 5% del numero di dati complessivi, corrispondente a 290 per ogni piattaforma. In seguito, è stata fatta un'ulteriore verifica al fine di assicurare che per ogni coppia dominio/mese-anno i valori non validi fossero inferiori a 10 per ogni parametro in quanto, in ultima istanza, i dati giornalieri sono stati aggregati attraverso operazioni di media ottenendo un dettaglio mensile. Questa operazione è stata effettuata attraverso una tabella pivot, motivo per cui i missing values e i valori non ammissibili rimasti non hanno avuto impatto sui dati aggregati. Il numero di siti rimasti dopo il trattamento ammonta a 111. Per il dataset sull'advertising con dettaglio mensile, sono stati eliminati tutti i siti non presenti nel dataset sull'engagement risultante dalle operazioni di pre-processing e i pochi valori pari a 0,00 rimasti sono stati sostituiti dalla media annuale dell'anno dell'osservazione del sito in esame in modo da non ridurre ulteriormente la numerosità del campione. Infine, per ottenere il dataset definitivo, i due set sono stati aggregati mantenendo unicamente come periodo di osservazione i mesi da settembre 2020 a settembre 2023. È stata a questo punto aggiunta un'ulteriore variabile, *OutgoingAD/Visits*, che indica il numero di click sui banner pubblicitari per visita. Tale parametro è considerato complementare all'*OutGoingAd Traffic* e isola l'effetto della policy sull'efficacia pubblicitaria rispetto all'andamento dell'engagement. Per tutti i siti risultanti da questa operazione e su cui non si aveva l'informazione sull'adozione della policy, sono state fatte ulteriori ricerche empiriche e sono stati eliminati tutti i domini per cui non è stato possibile ottenere un dettaglio completo

sull'implementazione, ottenendo così un totale di 100 domini con 37 osservazioni ciascuno.

3.3 Analisi preliminari

Prima di procedere con lo studio dei dati finalizzato a quantificare l'impatto della policy, risulta fondamentale eseguire delle analisi preliminari approfondite sul dataset a disposizione al fine di presentare una visione dettagliata delle caratteristiche fondamentali dei dati raccolti.

Calcolando la media annuale di tutte le variabili di engagement, è possibile notare un trend tendenzialmente negativo, fatta eccezione per la durata media delle visite e il rapporto pagine per visita, che subiscono una crescita dal 2020 al 2022 per poi diminuire nel 2023. Nota particolare per il Bounce Rate per cui un valore più basso è considerato migliore a livello di engagement a differenza delle altre variabili, che ha un trend negativo dal 2020 al 2022, per poi riprendere quota nel 2023. La variazione più significativa è osservabile nel totale di pagine visitate, che dal 2020 al 2023 rileva una diminuzione di quasi 100.000 pagine, sintomo della crisi discussa nei precedenti capitoli.

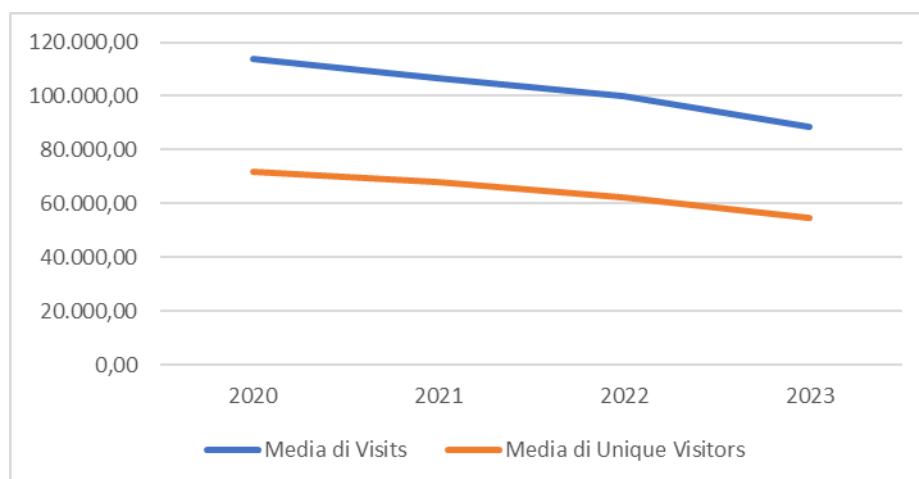


Fig. 19, Medie annuali dal 2020 al 2023 di Visite e Utenti Unici. Fonte: Elaborazione personale

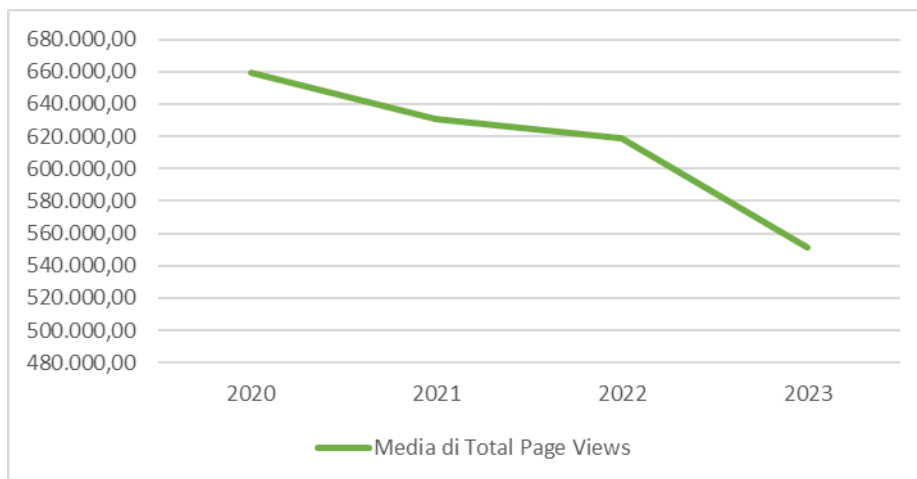


Fig. 20, Medie annuali dal 2020 al 2023 di Pagine Totali Visitate. Fonte: Elaborazione personale

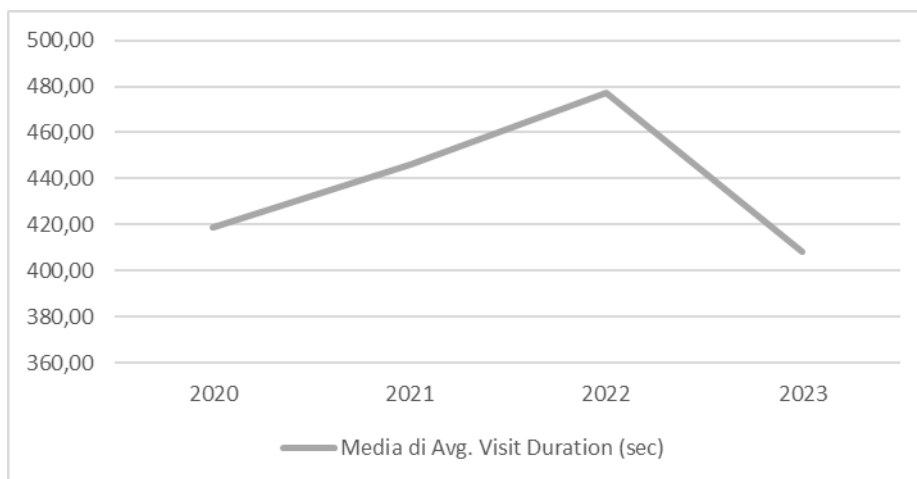


Fig. 21, Medie annuali dal 2020 al 2023 della Durata media delle visite in secondi. Fonte: Elaborazione personale

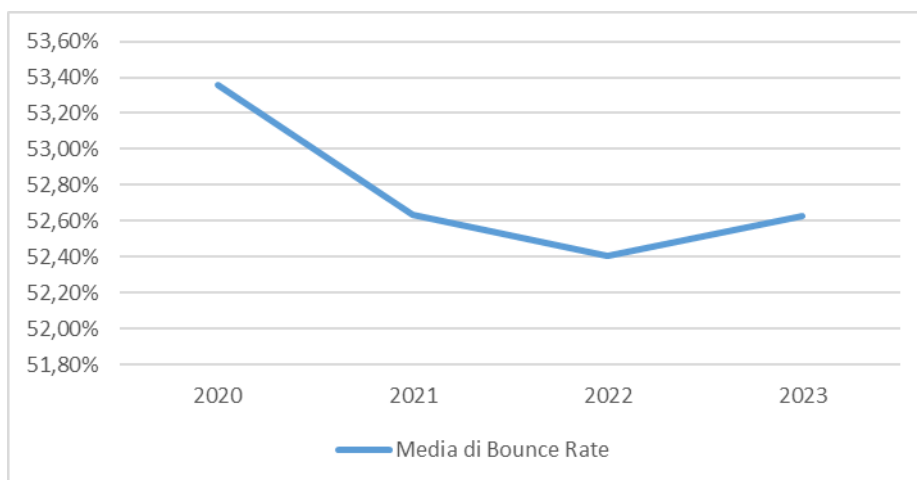


Fig. 22, Medie annuali dal 2020 al 2023 del Bounce Rate. Fonte: Elaborazione personale

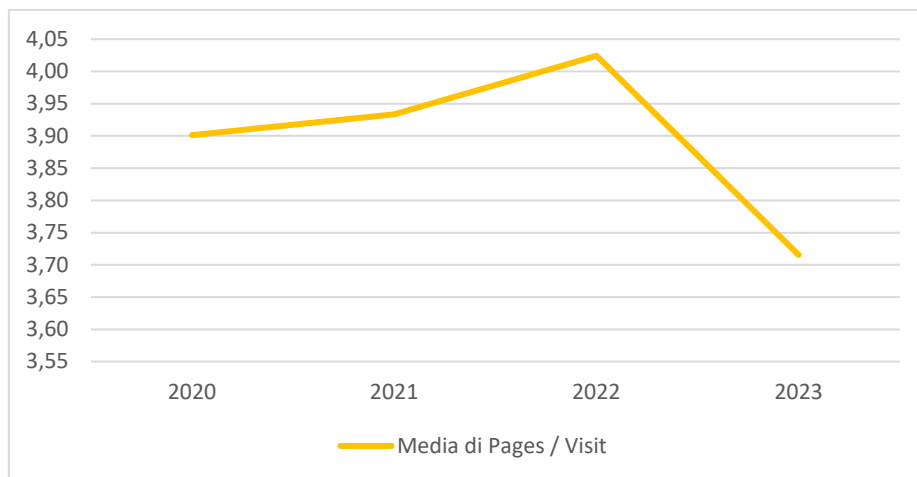


Fig. 23, Medie annuali dal 2020 al 2023 del rapporto Pages/Visit. Fonte: Elaborazione personale

Il trend tendenzialmente negativo dell'engagement si riflette anche sull'efficacia dell'advertising, che ha subito una forte diminuzione negli anni in esame.

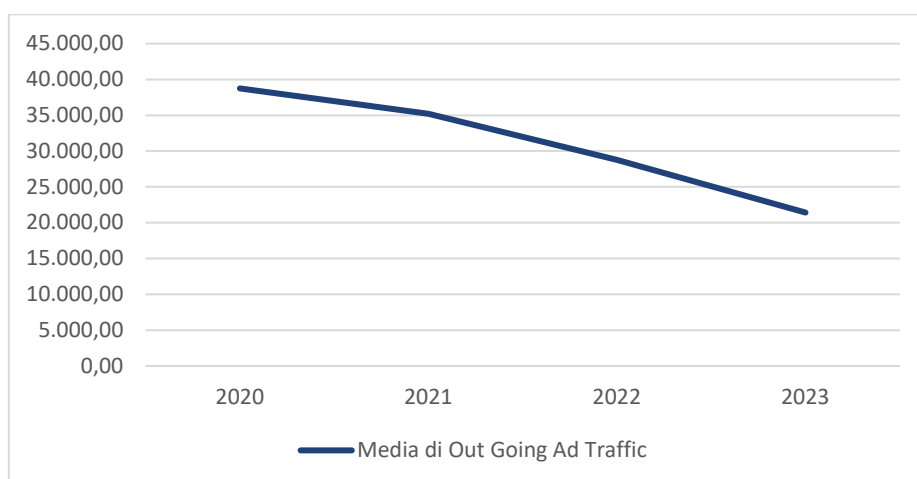


Fig. 24, Medie annuali dal 2020 al 2023 dell'Outgoing Ad Traffic. Fonte: Elaborazione personale

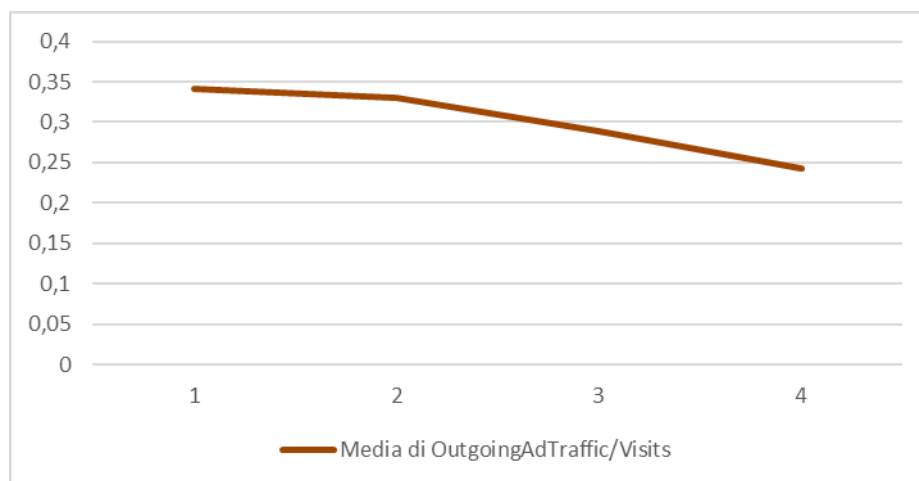


Fig. 25, Medie annuali dal 2020 al 2023 dell'Outgoing Ad Traffic. Fonte: Elaborazione personale

Dopo aver calcolato per ogni sito la media totale di ogni variabile, sono state calcolate le principali statistiche descrittive per ognuna di esse, riportate di seguito. In quasi tutte è osservabile un'alta variabilità tra i diversi siti, riscontrabile nei valori molto alti di deviazione standard e nell'elevata differenza tra valori di massimo e di minimo. Inoltre, la mediana posizionata molto prima della media nella maggioranza delle variabili, suggerisce che la distribuzione dei dati in questi casi è asimmetrica verso destra, ovvero che i valori molto elevati sono molti meno dei valori più bassi e che essi influenzano notevolmente la media, spingendola verso l'alto, mentre la mediana rimane relativamente bassa.

<i>Visits</i>		<i>Unique Visitors</i>		<i>Duration (sec)</i>		<i>Outgoing Ad Traffic</i>	
Media	100.756,26	Media	63.241,31	Media	444,02	Media	30.154,47
Mediana	24.034,65	Mediana	18.425,06	Mediana	369,41	Mediana	2.901,64
Deviazione standard	221.847,40	Deviazione standard	121.872,18	Deviazione standard	327,37	Deviazione standard	93.056,49
Minimo	8.567,39	Minimo	7.515,14	Minimo	62,98	Minimo	403,88
Massimo	1.337.752,88	Massimo	709.581,11	Massimo	1.471,11	Massimo	632.558,83
<i>Pages / Visit</i>		<i>Bounce Rate</i>		<i>Total Page Views</i>		<i>Media di OutgoingAd/Vists</i>	
Media	3,91	Media	52,64%	Media	610.591,58	Media	0,17
Mediana	3,54	Mediana	52,75%	Mediana	86.871,11	Mediana	0,13
Deviazione standard	2,05	Deviazione standard	0,13	Deviazione standard	1.878.400,71	Deviazione standard	0,15
Minimo	1,42	Minimo	14,59%	Minimo	16.450,80	Minimo	0,03
Massimo	12,09	Massimo	81,82%	Massimo	14.662.271,92	Massimo	1,21

Fig. 26, Statistiche descrittive. Fonte: Elaborazione personale

In seguito, è stata calcolata la matrice di correlazione. Osservando i dati, si può notare una forte correlazione tra le variabili *Visits*, *Unique Visitors*, *Total Page Views* e *OutGoing Ad Traffic*.

	Visits	Unique Visitors	Avg. Visit Duration (sec)	Pages / Visit	Bounce Rate	Total Page Views	Out Going Ad Traffic	OutgoingAdTraffic/Visits
Visits	1,00							
Unique Visitors	0,99	1,00						
Avg. Visit Duration (sec)	0,37	0,36	1,00					
Pages / Visit	0,42	0,39	0,63	1,00				
Bounce Rate	-0,37	-0,33	-0,48	-0,70	1,00			
Total Page Views	0,96	0,93	0,31	0,46	-0,41	1,00		
Out Going Ad Traffic	0,95	0,93	0,30	0,41	-0,38	0,95	1,00	
OutgoingAdTraffic/Visits	0,23	0,23	0,11	0,11	-0,16	0,21	0,33	1,00

Fig. 27, Matrice di correlazione. Fonte: Elaborazione personale

In conclusione, al fine di ottenere una stima preliminare dell'impatto del cookie wall sulle variabili in esame, è stato analizzato il comportamento di otto siti, di cui quattro che hanno implementato la policy (*lastampa.it*, *repubblica.it*, *corriere.it* e *ilfattoquotidiano.it*) e quattro che non l'hanno implementata (*ilpost.it*, *ilgiornale.it*, *liberoquotidiano.it* e *fanpage.it*), scegliendo tra tutti quelli che godono di una più vasta popolarità. Per ognuno dei due gruppi sono state calcolate le medie per tutti i mesi di osservazione per ogni parametro, successivamente normalizzate, al fine di eliminare le differenze di livello tra i due campioni ed ottenere un'analisi più chiara. Di seguito i grafici risultanti per ogni variabile, in cui sono presenti le serie temporali normalizzate dei due gruppi ed è evidenziato il momento in cui la policy è entrata in vigore, ovvero durante il mese di ottobre 2022 per il campione considerato. Si rimanda all'appendice per il dettaglio su ogni sito preso in considerazione in questa analisi preliminare.

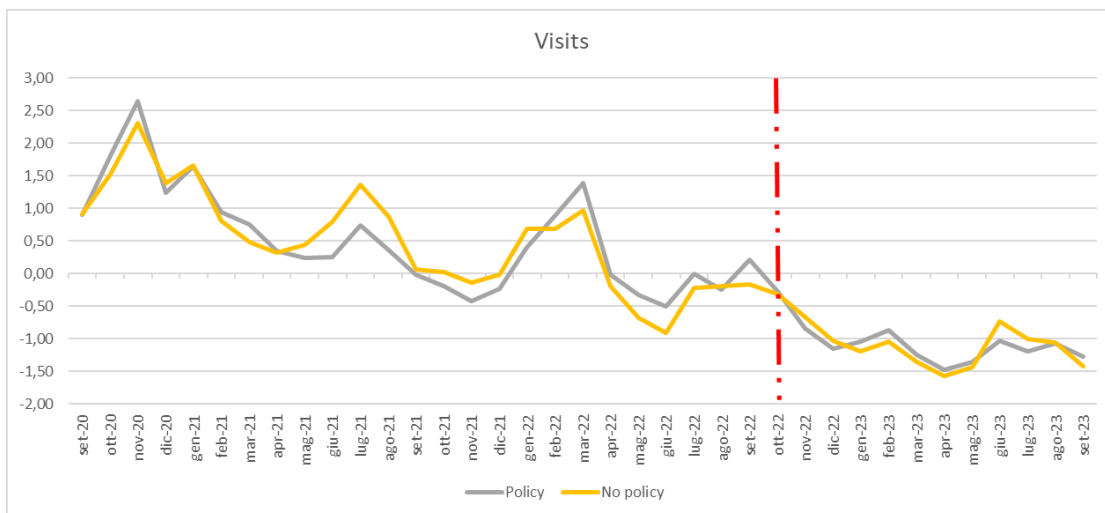


Fig. 28, Serie temporali Visits Fonte: Elaborazione personale

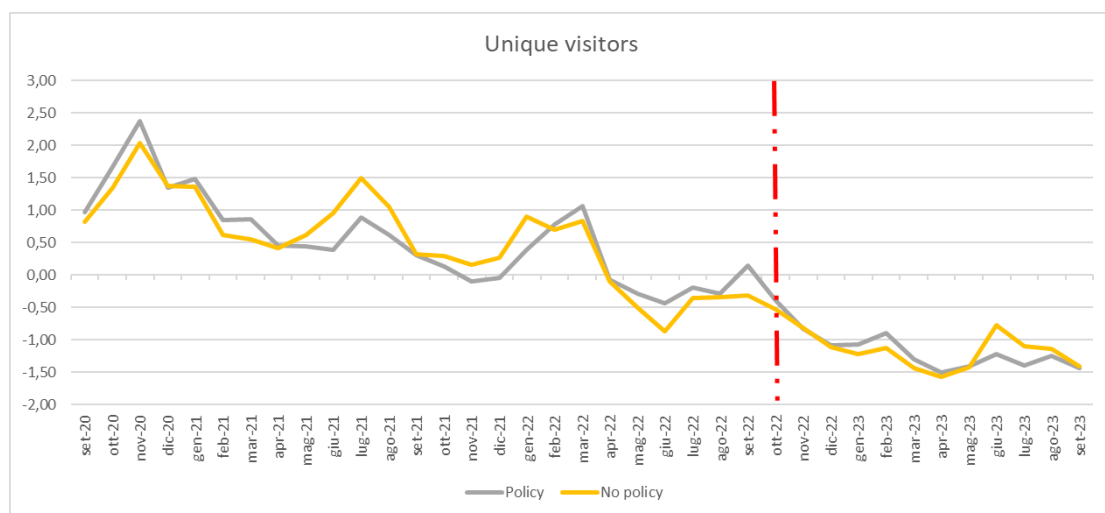


Fig. 29, Serie temporali Unique Visitors. Fonte: Elaborazione personale

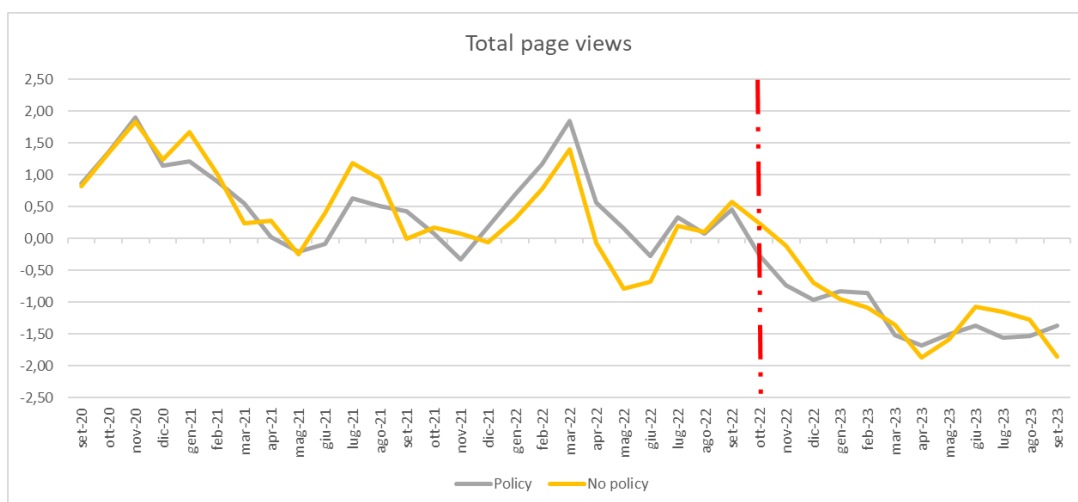


Fig. 30, Serie temporali Total page views. Fonte: Elaborazione personale

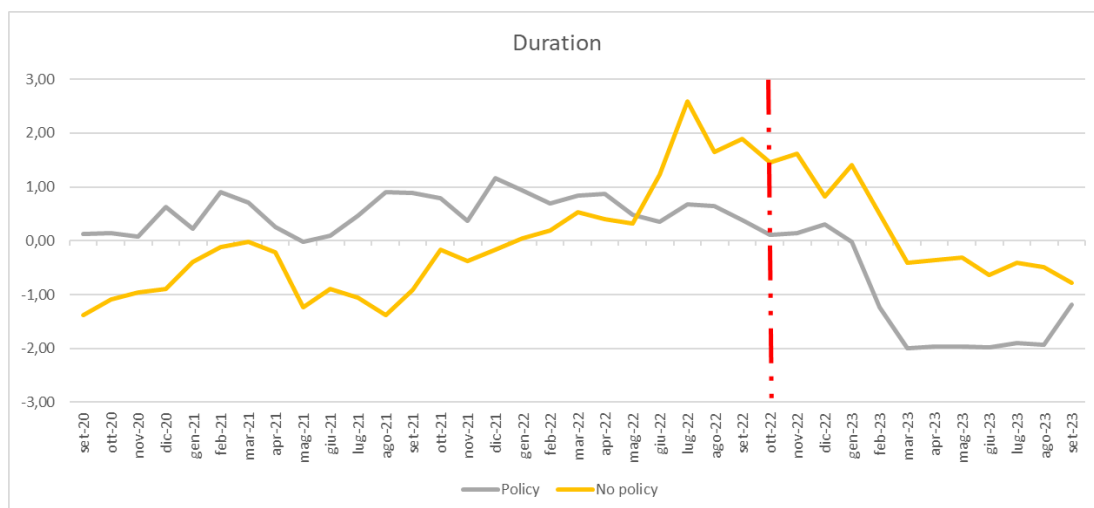


Fig. 31, Serie temporali Visit duration. Fonte: Elaborazione personale

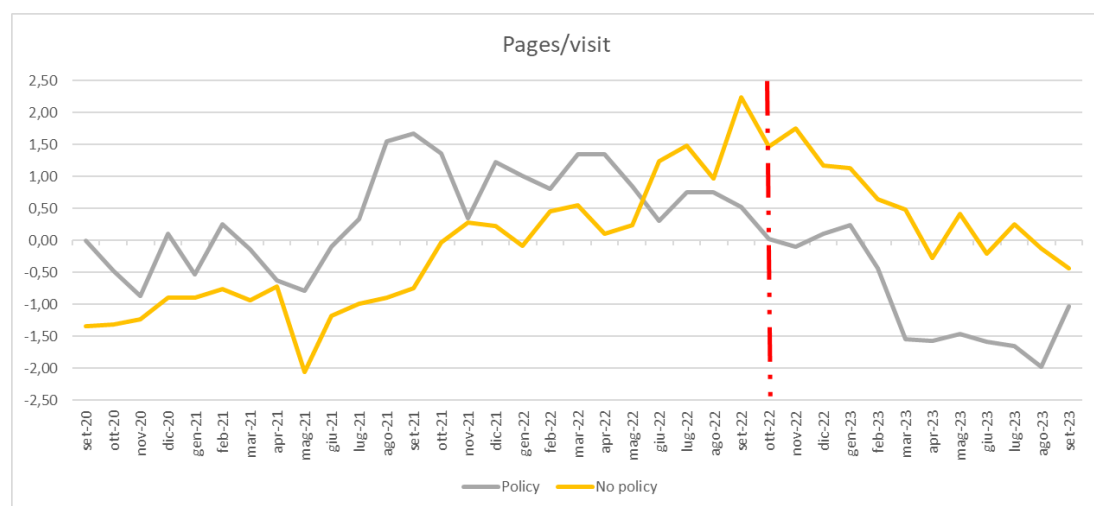


Fig. 32, Serie temporali Pages/visit. Fonte: Elaborazione personale

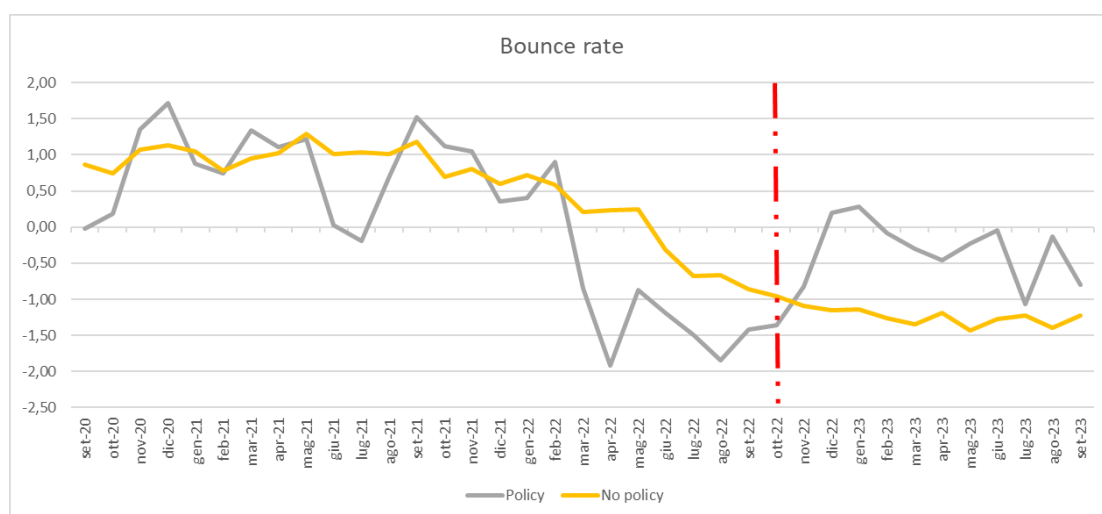


Fig. 33, Serie temporali Bounce rate. Fonte: Elaborazione personale

Analizzando l'engagement, per le variabili correlate *Visits*, *Unique Visitors* e *Total Page Views* è evidente un andamento temporale negativo per entrambi i campioni come preannunciato dalle precedenti analisi e, in corrispondenza dell'implementazione della policy, non è possibile notare un significativo cambiamento di trend per nessuno dei due gruppi. Per i parametri *Duration* e *Pages/Visit* è possibile osservare un trend differente per i due gruppi: per il primo periodo l'andamento è tendenzialmente in crescita per entrambi con un vantaggio per il gruppo che ha implementato la policy, fino ai mesi di maggio e giugno del 2022, in cui il gruppo che non ha implementato la policy continua a crescere anche più rapidamente di prima, acquistando il dominio rispetto all'altro, il quale inizia ad avere una tendenza negativa. Successivamente a tale periodo, entrambi hanno un trend decrescente, ma in ogni caso continua ad essere ben evidente il vantaggio del gruppo che non ha implementato la policy. Vista tale evidenza, si sono condotte ulteriori ricerche finalizzate a capire se nel periodo di cambio trend c'è stato effettivamente uno shock per il mercato, ma non sono state trovate evidenze rilevanti. Per la variabile *Bounce rate*, per la quale un valore più alto è sinonimo di un engagement più insoddisfacente, il trend è tendenzialmente negativo e, in prossimità dell'implementazione della policy, si osserva un'inversione tra i due gruppi, passando da una situazione tendenzialmente più vantaggiosa per i siti che hanno implementato la policy all'esatto opposto.

Alla luce delle precedenti osservazioni, è possibile presumere che la policy abbia avuto un effetto negativo sull'engagement dei siti che l'hanno implementata. Tuttavia, queste analisi preliminari, non risultano sufficienti a confermare questa tesi con un buon livello di sicurezza. Si procede ora con l'analisi sull'advertising.

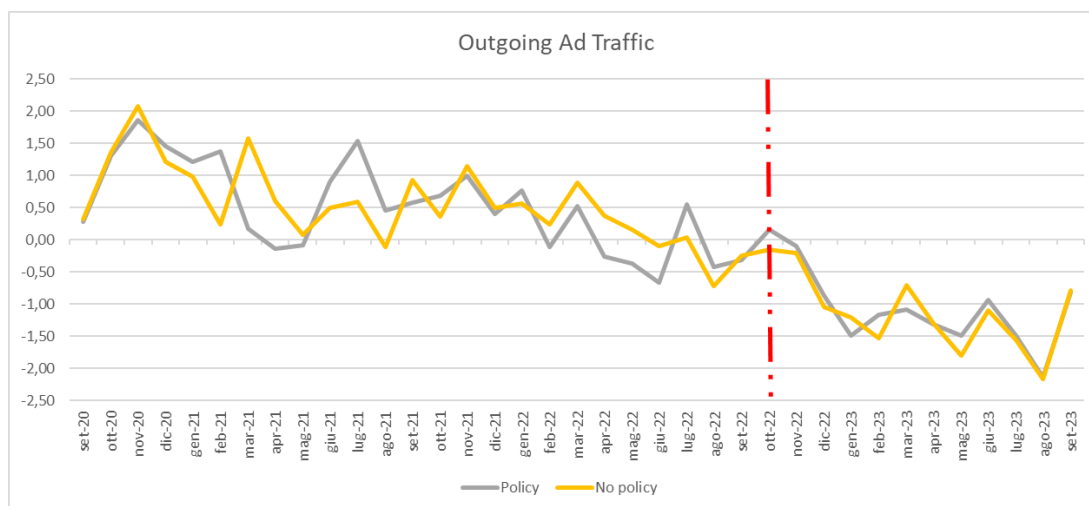


Fig. 34, Serie temporali Outgoing Ad Traffic. Fonte: Elaborazione personale

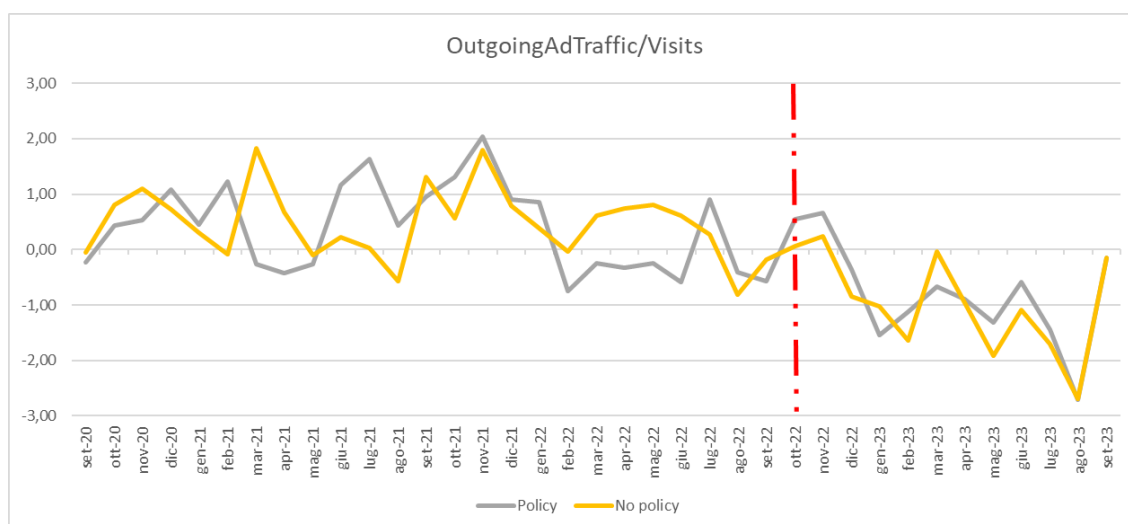


Fig. 35, Serie temporali OutgoingAdTraffic/Visits. Fonte: Elaborazione personale

Guardando l'andamento di entrambe le variabili, l'efficacia dell'advertising, presenta un trend negativo per tutto il periodo in esame. Al contrario delle aspettative inoltre, il cookie wall sembra non migliorare in modo significativo le performance della pubblicità, nonostante permetta una profilazione più accurata degli utenti. Anche in questo caso però l'analisi preliminare, non risulta sufficiente a confermare questa tesi con un buon livello di sicurezza.

3.4 Analisi econometrica

Al fine di individuare ed analizzare gli effetti quantitativi del cookie wall sull'engagement e sull'advertising, si procederà con un'analisi econometrica, in particolare con l'implementazione di modelli *Diff-in-Diff*. Il metodo "*Difference-in-Differences*" (Diff-in-Diff) rappresenta una metodologia utilizzata nell'analisi degli effetti causali di un intervento o di un trattamento su un gruppo di osservazioni nel tempo. Questo approccio si basa sull'idea di confrontare le variazioni nel tempo tra un gruppo di trattamento, ovvero il gruppo che ha subito l'intervento, e un gruppo di controllo, ovvero il gruppo che non ha subito l'intervento, misurando la differenza nella differenza tra i due. Il diff-in-diff, per individuare l'impatto di un determinato trattamento, misura quindi due differenze:

- 1) Differenza tra il pre e il post trattamento all'interno di ciascun gruppo;
- 2) Differenza tra i due gruppi.

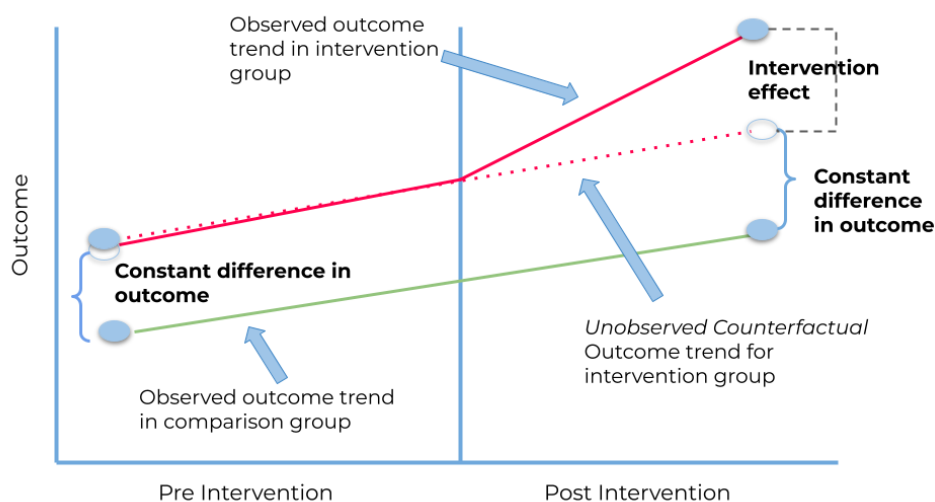


Fig. 36, Diff-in-diff. Fonte: Elaborazione personale

Nel caso in esame, il gruppo di trattamento è rappresentato dai domini che hanno adottato il cookie wall, mentre il gruppo di controllo è rappresentato dai

domini che non lo hanno adottato. I giornali del gruppo CEDS che hanno adottato la variante della policy analizzata nel paragrafo 3.2 sono stati inseriti nel gruppo di controllo, in quanto la forma alternativa da loro adottata non obbliga il consumatore non abbonato ad accettare i cookie di profilazione per accedere ai contenuti.

Dato che la policy in esame non ha un'unica data di adozione per tutti i siti del gruppo di trattamento, in questa sede verrà utilizzato un Diff-in-Diff con *staggered adoption*, che consente di gestire l'implementazione non uniforme nel tempo. In particolare, per poter generare i modelli, al dataset sopra descritto è stata aggiunta una dummy nominata *PrivacyWall*, in grado di catturare sia l'informazione sull'appartenenza al gruppo di trattamento o al gruppo di controllo, sia l'informazione temporale sull'implementazione della policy in quanto assume:

- 1) valore 0 per tutte le osservazioni del gruppo di controllo e per tutte le osservazioni del gruppo di trattamento antecedenti alla data di implementazione della policy;
- 2) valore 1 per tutte le osservazioni del gruppo di trattamento successive alla data di adozione del cookie wall.

Per stimare l'impatto verranno utilizzati i seguenti modelli:

$$1) \log(Y_{i,t}) = \beta_0 + \beta_1 \text{PrivacyWall}_{i,t} + FE_i + FE_t + \varepsilon_{i,t}$$

$$2) Y_{i,t} = \beta_0 + \beta_1 \text{PrivacyWall}_{i,t} + FE_i + FE_t + \varepsilon_{i,t}$$

La regressione 1 verrà applicata alle variabili *Visits*, *Average Visit Duration*, *Pages/Visit* e *Outgoing Ad Traffic* per le quali verrà studiato il logaritmo in modo da poter interpretare i coefficienti in termini percentuali, mentre la regressione 2 sarà applicata alle variabili *Bounce rate* e *OutGoingAdTraffic/Visits*, per cui una variazione lineare è considerata più chiara. Le variabili *Unique Visitors* e *Total*

Page Views sono state escluse dall'analisi in quanto strettamente correlate con il parametro *Visits* (coefficiente di correlazione $> 0,90$) e per cui un ulteriore studio sarebbe risultato ridondante. In entrambi i modelli β_1 cattura l'impatto casuale del cookie wall sull'outcome di interesse e sono stati inseriti effetti fissi a livello di dominio ed effetti fissi a livello di tempo, i quali controllano per i fattori idiosincratici non osservabili ma presumibilmente importanti per descrivere la variazione della variabile dipendente. In particolare gli effetti fissi sul sito controllano per i fattori specifici del sito che non variano nel tempo e che spiegano la differenza tra livelli differenti delle diverse variabili per i domini in esame, mentre gli effetti fissi sul tempo controllano per i diversi trend temporali. Inoltre, nelle regressioni sono stati clusterizzati gli errori standard in base all'ID, ovvero un codice alfa-numerico che identifica univocamente ogni dominio, tenendo conto della possibilità che le osservazioni all'interno dello stesso cluster possano essere correlate in modo più stretto rispetto a quelle in altri cluster e al fine di ottenere stime più efficienti dei coefficienti del modello. Nelle tabelle dei risultati che seguono, i coefficienti della variabile indipendente Privacywall delle colonne 1, 2, 3 e 5 indicano come varia in termini percentuali la variabile dipendente a cui sono riferiti quando è rilevata la presenza del cookie wall, mentre per le colonne 4 e 5 la stessa variabilità è espressa in termini lineari. Sotto ogni coefficiente è indicato in parentesi il relativo errore standard. Sono presenti inoltre il numero totale di osservazioni N e R^2 , il quale indica la percentuale della variabilità della variabile dipendente spiegata dal modello.

3.4.1 Diff-in-Diff primo caso

Nel caso in esame, il dataset con livello di dettaglio mensile è stato impiegato interamente, contando quindi 9 domini nel gruppo di trattamento e 91 nel gruppo di controllo e 37 osservazioni per ogni sito corrispondenti ai mesi da

settembre 2020 a settembre 2023. Di seguito i risultati ottenuti applicando i modelli descritti nel paragrafo precedente.

	(1) log Visits	(2) log Duration	(3) log PagesVisit	(4) BounceRate	(5) log AD	(6) OutGoingAdVisits
Privacywall	-0.0884 (0.0926)	-0.223 (0.180)	-0.0974 (0.0781)	0.00752 (0.0116)	-0.0918 (0.156)	-0.0538** (0.0264)
<i>N</i>	3700	3700	3700	3700	3700	3700
<i>R</i> ²	0.972	0.871	0.873	0.866	0.858	0.599

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Tab.1, Risultati Diff-in-Diff dataset completo con dettaglio mensile. Fonte: Stata

È interessante notare che i coefficienti ottenuti mostrino effettivamente un impatto negativo del cookie wall sulle variabili di engagement, in linea con le aspettative. Tuttavia, tali risultati non hanno raggiunto la significatività statistica e quindi, nonostante l'apparente associazione tra l'engagement e il cookie wall, non è possibile considerare tali relazioni come affidabili o indicative di un effetto significativo. Per quanto riguarda invece le due variabili sull'Advertising, in collisione con l'obiettivo dell'implementazione della policy, l'efficacia dello stesso sembra diminuire. In particolare per la variabile in colonna 4 non si è raggiunta la significatività statistica, per cui non è possibile considerare rilevante l'effetto espresso dal coefficiente, mentre la colonna 5 raggiunge una significatività del 5%, indicando che il rapporto tra i click sui banner pubblicitari e le visite totali diminuisce in media di circa 0.05 in presenza del cookie wall.

3.4.2 Diff-in-Diff secondo caso

Alla luce dei risultati precedenti, al fine di ottenere un'ulteriore evidenza, è stato ridotto il dataset di partenza in modo da offrire al gruppo di trattamento un miglior controfattuale, in quanto, la maggior parte dei siti che hanno adottato la policy di interesse, risultano essere di particolare rilievo e con livelli di engagement considerevoli, mentre, nel dataset integrale, tra i siti che fanno

parte del gruppo di controllo è presente un'elevata variabilità sotto questo punto di vista. Per ridimensionare il dataset è stata calcolata la quota di mercato di ciascun dominio utilizzando il parametro *Visits* con riferimento all'anno che va da settembre 2020 a settembre 2021, periodo scelto poiché sufficientemente lontano dall'implementazione della policy e quindi non influenzato dalla stessa. Il mercato è risultato particolarmente concentrato, raggiungendo un livello di copertura dell'85% con i primi 32 domini, quelli poi effettivamente considerati nel dataset ridotto in modo da garantire una significatività allo studio.

<i>Dominio</i>	<i>Quota di mercato cumulata</i>	<i>Dominio</i>	<i>Quota di mercato cumulata</i>	<i>Dominio</i>	<i>Quota di mercato cumulata</i>	<i>Dominio</i>	<i>Quota di mercato cumulata</i>
repubblica.it	0,176	liberoquotidiano.it	0,592	dilei.it	0,732	leggo.it	0,803
corriere.it	0,301	dagospia.com	0,616	rainews.it	0,742	iltempo.it	0,811
ansa.it	0,379	gelocal.it	0,638	ilrestodelcarlino.it	0,752	unionesarda.it	0,817
ilfattoquotidiano.it	0,435	ilmessaggero.it	0,659	adnkronos.com	0,762	lanazione.it	0,824
news.google.com	0,483	ilpost.it	0,678	today.it	0,771	ilgiorno.it	0,830
lastampa.it	0,513	notizie.it	0,695	ilmattino.it	0,780	romatoday.it	0,836
fanpage.it	0,540	huffingtonpost.it	0,708	ilgazzettino.it	0,788	ilsecoloxix.it	0,841
ilgiornale.it	0,567	quotidiano.net	0,720	open.online	0,796	milanotoday.it	0,845

Fig. 37, Quote di mercato cumulate. Fonte: Elaborazione personale

A supporto della tesi sopracitata, nelle prime quattro posizioni, raggiungendo una quota di mercato cumulata pari al 44%, si posizionano tutti siti che hanno adottato la policy. Inoltre, dei 9 giornali iniziali appartenenti al gruppo di trattamento, 6 rientrano nelle prime 32 posizioni, confermando ancora una volta come il cookie wall abbia interessato prevalentemente i big player del mercato. Di seguito i risultati ottenuti.

	(1) log Visits	(2) log Duration	(3) log PagesVisit	(4) BounceRate	(5) log AD	(6) OutGoingAdVisits
Privacywall	-0.0843 (0.126)	-0.435* (0.246)	-0.188* (0.110)	0.0553*** (0.0164)	0.144 (0.226)	-0.0189 (0.0325)
<i>N</i>	1184	1184	1184	1184	1184	1184
<i>R</i> ²	0.939	0.891	0.875	0.896	0.849	0.541

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Tab.2, Risultati Diff-in-Diff dataset ridotto con dettaglio mensile. Fonte: Stata

In questa seconda analisi, sono presenti degli standard error sicuramente più alti vista la diminuzione della numerosità del campione. I coefficienti legati alle

variabili di engagement mostrano un peggioramento in questo senso dovuto all'implementazione della policy maggiore rispetto al caso precedente. Al contrario dei precedenti risultati, in questo caso tre coefficienti su quattro hanno raggiunto la significatività statistica. In particolare la *Duration* e il rapporto *Pages/Visit* hanno raggiunto una significatività del 10% e mostrano rispettivamente una diminuzione in media del 43.5% e del 18.8% in presenza del cookie wall, mentre il *Bounce Rate* ha raggiunto una significatività del 1% e mostra un aumento in media di 0.053 punti percentuali dopo l'implementazione della policy. Per quanto riguarda l'Advertising, la variabile *OutgoingAdTraffic* sembra avere un miglioramento con l'implementazione della policy, miglioramento che viene però a mancare se rapportiamo la stessa al numero di visite. Tuttavia, nessuno dei due coefficienti raggiunge la significatività statistica, motivo per cui in questo caso non risulta un impatto significativo sull'efficacia della pubblicità.

3.4.3 Diff-in-Diff terzo caso

Con l'obiettivo di studiare gli effetti che ha avuto il cookie wall nel breve periodo, in concomitanza con la sua implementazione, è stato costruito un ulteriore dataset con livello di dettaglio giornaliero, prendendo come riferimento i 100 giornali protagonisti anche delle analisi iniziali e come periodo di riferimento quello che va da settembre 2022 ad aprile 2023, in quanto i siti del gruppo di trattamento hanno implementato la policy tra il 17 ottobre 2022 e il 7 Marzo 2023. L'analisi è rivolta unicamente alle variabili di engagement, in quanto per le variabili di advertising non è disponibile il dettaglio giornaliero. Di seguito i risultati ottenuti.

	(1) log_Visits	(2) log_Duration	(3) log_PagesVisit	(4) BounceRate
Privacywall	-0.0390 (0.0514)	-0.182* (0.103)	-0.0677 (0.0521)	0.0000703 (0.00940)
<i>N</i>	24200	24200	24200	24200
<i>R</i> ²	0.976	0.738	0.816	0.819

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Tab.3, Risultati Diff-in-Diff dataset completo con dettaglio giornaliero. Fonte: Stata

Anche in questo caso l'engagement mostra un declino in presenza del cookie wall, seppur con una portata minore rispetto alle analisi precedenti. L'unico risultato che raggiunge la significatività statistica al 10%, e che può quindi essere considerato attendibile, è la diminuzione in media del 18.2% della *Duration* con l'implementazione della policy.

3.4.4 Diff-in-Diff quarto caso

Anche per il dataset con dettaglio giornaliero è stato studiato il campione più ridotto, composto dai 32 siti illustrati nel paragrafo 3.4.2. Di seguito i risultati ottenuti.

	(1) log_Visits	(2) log_Duration	(3) log_PagesVisit	(4) BounceRate
Privacywall	-0.102** (0.0450)	-0.274* (0.138)	-0.112 (0.0681)	0.0165* (0.00845)
<i>N</i>	7744	7744	7744	7744
<i>R</i> ²	0.984	0.825	0.867	0.868

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Tab.4, Risultati Diff-in-Diff dataset ridotto con dettaglio giornaliero. Fonte: Stata

Si riconferma nuovamente il calo dell'engagement, in questo caso con tre coefficienti significativi: la *Duration* e il *Bounce Rate* che raggiungono una significatività del 10% e mostrano rispettivamente una diminuzione del 27.4% in media e un aumento di 0.0165 punti percentuali, e le *Visits*, su cui osserviamo per la prima volta un calo significativo al 5% del 10.2% in media.

CONCLUSIONI

I siti di informazione online sono un tipico esempio di piattaforma digitale che opera in un mercato a due versanti. Da un lato, gli utenti consumano contenuti informativi, mentre, dall'altro, le aziende pubblicitarie cercano di raggiungere questa base di utenti. Questo crea una dinamica interconnessa dove la presenza e l'interazione dei visitatori influenzano direttamente il valore offerto agli inserzionisti e viceversa. Gli effetti di rete cross side giocano quindi un ruolo chiave, motivo per il quale la struttura dei prezzi è spesso non neutrale e può includere modelli come l'offerta di contenuti gratuiti finanziati interamente o quasi dalla pubblicità. Uno degli aspetti centrali in questo contesto è la raccolta e l'analisi dei dati sugli utenti, le quali consentono una pubblicità più mirata, migliorandone l'efficacia. Questa capacità di offrire agli inserzionisti un pubblico targettizzato e più incline ai loro prodotti o servizi può rendere la pubblicità più profittevole, consentendo alle piattaforme di richiedere tariffe più elevate. È in questo contesto che tra la fine del 2022 e l'inizio del 2023 i più grandi gruppi editoriali italiani, al fine di migliorare il trend negativo dei ricavi pubblicitari degli ultimi anni, hanno deciso di implementare sulle loro pagine web il cookie wall, adottando una policy che ingloba la strategia del *pay or consent* già adottata dai colleghi esteri qualche tempo prima. Tale policy mette gli utenti interessati ad usufruire del contenuto del sito davanti a due opzioni: abbonarsi al servizio o prestare il consenso per i cookie di profilazione, offrendo alla piattaforma dati preziosi per la targhettizzazione. Il presente studio ha valutato l'impatto di tale strategia sull'engagement e sull'advertising per le piattaforme che hanno deciso di adottarla. L'analisi è stata condotta attraverso l'utilizzo di modelli Diff-in-Diff, adatti allo studio degli effetti causali di un intervento o di un trattamento. Si sottolinea come la policy in oggetto non sia completamente esogena in quanto sono gli stessi siti a decidere se

implementarla o meno, caratteristica che potrebbe avere degli effetti sui risultati. Per l'analisi sono stati presi in considerazione quattro dataset diversi:

- 1) Dataset completo composto da 100 giornali con livello di dettaglio mensile con periodo di osservazione da settembre 2020 a settembre 2023;
- 2) Dataset ridotto composto da 32 giornali con livello di dettaglio mensile con periodo di osservazione da settembre 2020 a settembre 2023;
- 3) Dataset completo composto da 100 giornali con livello di dettaglio giornaliero con periodo di osservazione da settembre 2022 ad aprile 2023 solo per le variabili di engagement;
- 4) Dataset ridotto composto da 32 giornali con livello di dettaglio giornaliero con periodo di osservazione da settembre 2022 ad aprile 2023 solo per le variabili di engagement.

Il dataset ridotto prende in considerazione solo i big player del mercato delle news online, considerato comunque abbastanza concentrato, tanto da raggiungere una copertura dell'85% considerando solo i primi 32 siti online per numero di visite. Riducendo il dataset, si è infatti proposto un controllo più credibile in quanto più vicino al gruppo di trattamento. Difatti, come si evince dalla tesi, il trattamento non è da considerare randomizzato in quanto sono solo le grandi testate, o comunque i grandi gruppi editoriali, ad aver scelto di implementare la policy. I consumatori vengono considerati inelastici al cookie wall nel momento in cui la scelta è tra uno dei big player e un sito meno prestigioso, come dimostrano tendenzialmente i risultati delle analisi che utilizzano il dataset completo, in cui infatti la policy, in generale, non mostra un impatto significativo a livello di engagement. Studiando il dataset ridotto e avendo quindi un miglior controfattuale, i risultati più evidenti sul lungo periodo sono riferiti all'*intensive margin*, ovvero l'uso intensivo una volta che l'utente ha aperto il sito, mentre sull'*extensive margin*, ovvero il numero di utenti

che approcciano al sito, la policy non ha avuto un impatto rilevante, a meno del primo periodo in cui si è registrata una diminuzione significativa del numero di visite prendendo in considerazione i dati giornalieri dei big player. Questo potrebbe essere sintomo del fatto che gli utenti in qualche modo, nel momento in cui si ritrovano un banner che potrebbe essere considerato invasivo e che costringe ad accettare la loro profilazione, si tirano indietro, interagendo in maniera minore con i siti che hanno implementato la policy, preferendo invece quelli che permettono una libera fruizione dei contenuti, se altrettanto validi e di spessore. Andando più nel dettaglio, l'impatto negativo subito dall'*intensive margin* potrebbe anche dipendere dal fatto che gli utenti già registrati o abituali potrebbero percepire il cookie wall come pressante, mettendo in discussione la gestione della privacy e reagendo negativamente alla richiesta immediata di accettare i cookie di profilazione o sottoscrivere un abbonamento. Essi infatti, avendo già sviluppato una relazione con il sito, potrebbero resistere a cambiamenti improvvisi nell'esperienza utente. D'altra parte, l'*extensive margin*, potrebbe aver retto al cookie wall grazie ai nuovi visitatori, i quali tendenzialmente potrebbero mostrare una maggiore tolleranza. Questi utenti, ancora in fase di esplorazione e privi di legami preesistenti con il sito, potrebbero essere meno influenzati dalla richiesta di consenso iniziale, focalizzandosi sull'accesso immediato ai contenuti. La loro percezione potrebbe essere meno sensibile ai cambiamenti nelle politiche di privacy, poiché non hanno ancora sviluppato una conoscenza approfondita del sito.

Per quanto riguarda invece l'advertising da quanto si evince dai risultati, l'implementazione del un cookie wall, al contrario delle aspettative, non ha migliorato l'efficacia della pubblicità, anzi in alcuni casi sembra averla peggiorata. Una delle ipotesi dietro questo mancato miglioramento potrebbe essere la riluttanza degli utenti nel concedere il consenso tramite il cookie wall,

tesi supportata anche dalla riduzione dell'engagement provata dalle analisi condotte, limitando così la raccolta dati. Inoltre, se il cookie wall è stato interpretato dai visitatori come un'invasione nella loro privacy, potrebbe aver minato la fiducia necessaria per un coinvolgimento efficace con gli annunci pubblicitari. Paradossalmente, un banner meno d'impatto e meno vincolante, come quello presente nella maggioranza dei siti che non hanno adottato i cookie wall, potrebbe recare meno sfiducia nell'utente che accetta in modo più leggero l'opzione della profilazione.

In conclusione, dal presente studio emerge chiaramente che gli utenti non sono indifferenti alle strategie implementate dai siti online, specialmente in relazione alla raccolta di dati e alle questioni legate alla privacy. La riluttanza degli utenti nel concedere il consenso attraverso meccanismi come il cookie wall suggerisce che l'attenzione alla privacy svolge un ruolo cruciale nelle dinamiche di engagement e nell'efficacia delle strategie pubblicitarie. La fiducia degli utenti e l'adozione di meccanismi che consentono loro di esercitare un controllo più diretto sulla propria esperienza, è fondamentale per migliorare l'engagement e l'efficacia pubblicitaria e mantenere relazioni positive, e, qualsiasi approccio che rischia di essere percepito come invasivo, rischia di compromettere l'interazione positiva con il pubblico. In un panorama digitale in costante evoluzione, la considerazione attenta delle esigenze degli utenti è imprescindibile per il successo di qualsiasi strategia online.

APPENDICI

APPENDICE A

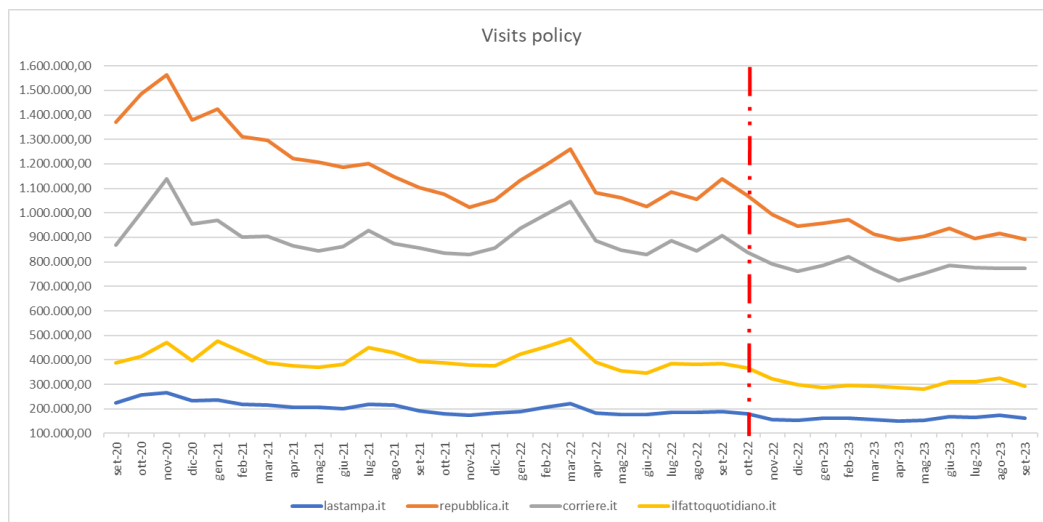


Fig. A.1, Serie temporali Visits giornali policy. Fonte: Elaborazione personale

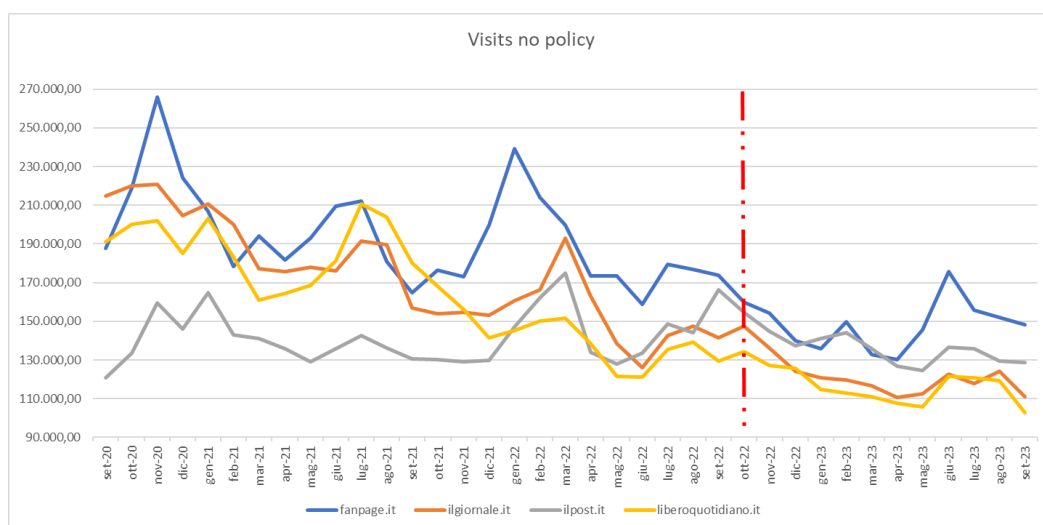


Fig. A.2, Serie temporali Visits giornali no policy. Fonte: Elaborazione personale

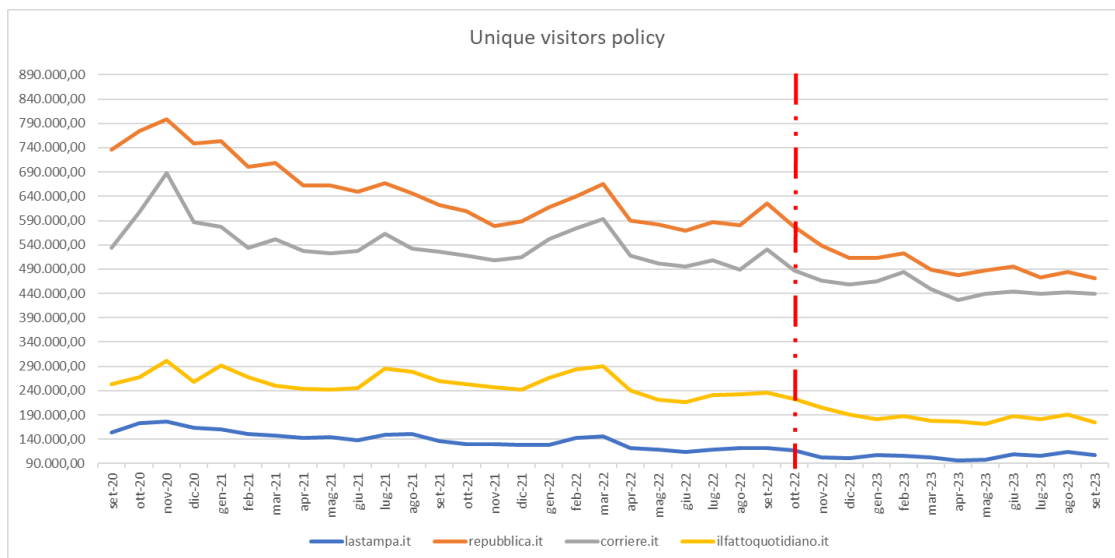


Fig. A.3, Serie temporali Unique visitors giornali policy. Fonte: Elaborazione personale

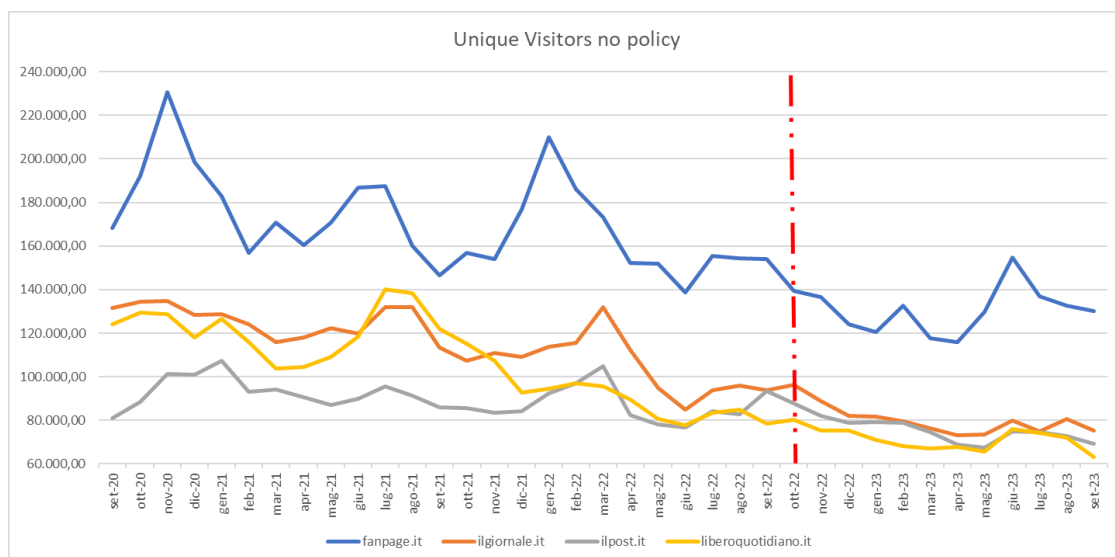


Fig. A.4, Serie temporali Unique visitors giornali no policy. Fonte: Elaborazione personale

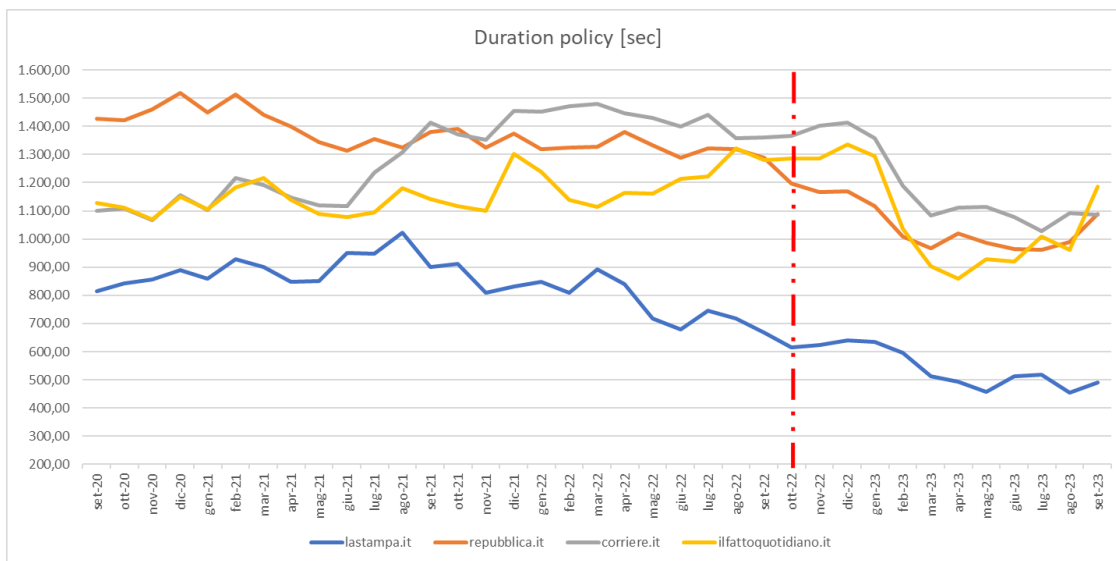


Fig. A.5, Serie temporali Duration giornali policy. Fonte: Elaborazione personale

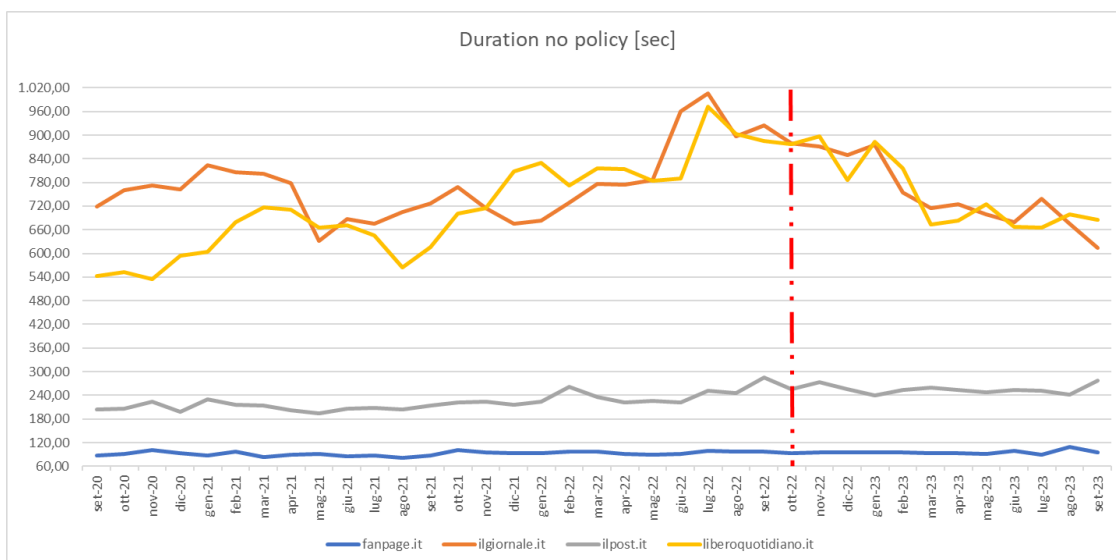


Fig. A.6, Serie temporali Duration giornali no policy. Fonte: Elaborazione personale

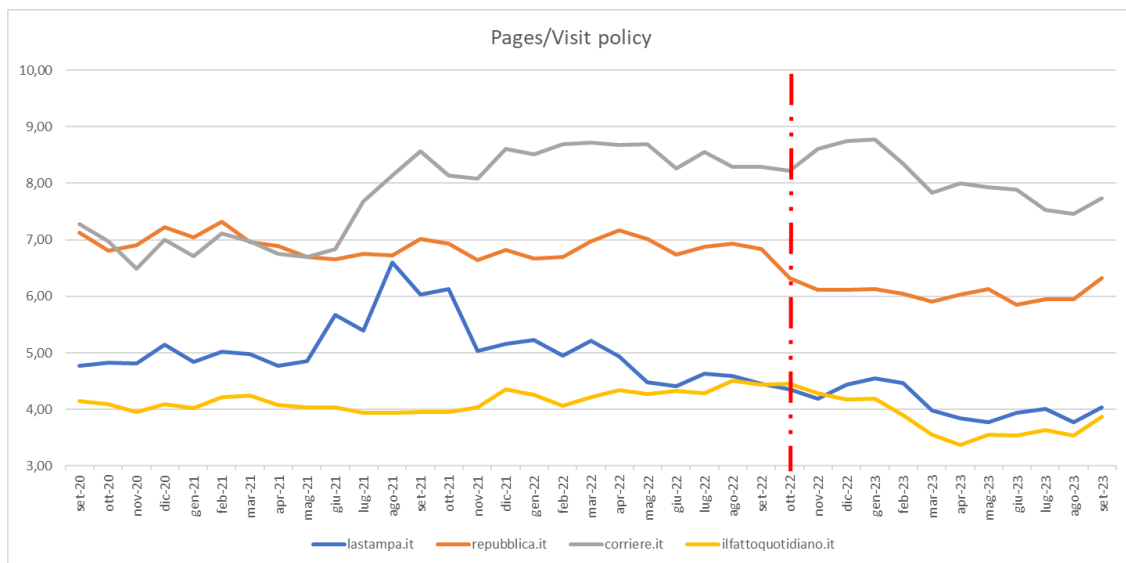


Fig. A.7, Serie temporali Pages/Visit giornali policy. Fonte: Elaborazione personale

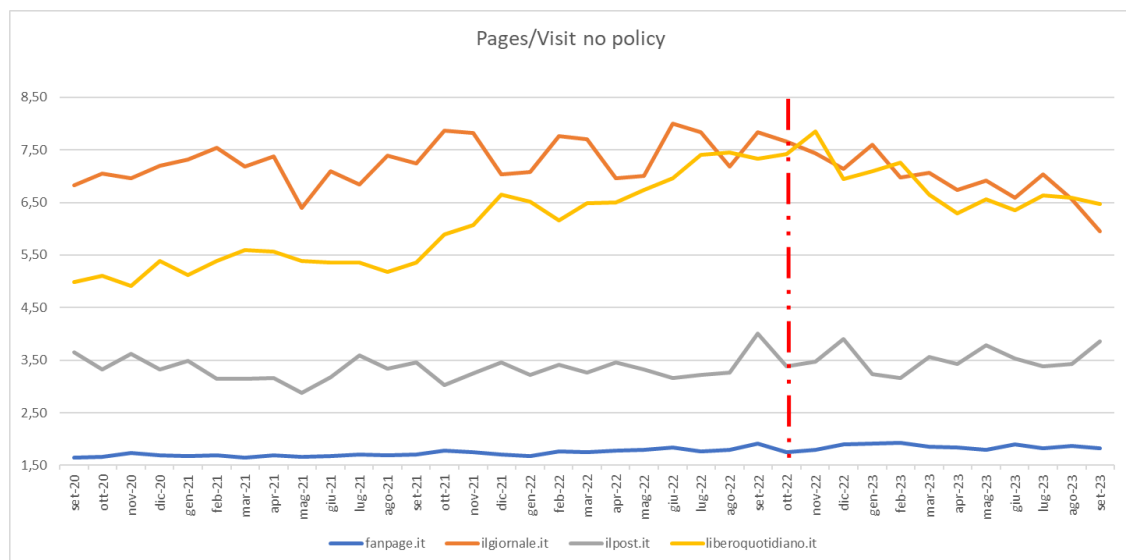


Fig. A.8, Serie temporali Pages/Visit giornali no policy. Fonte: Elaborazione personale

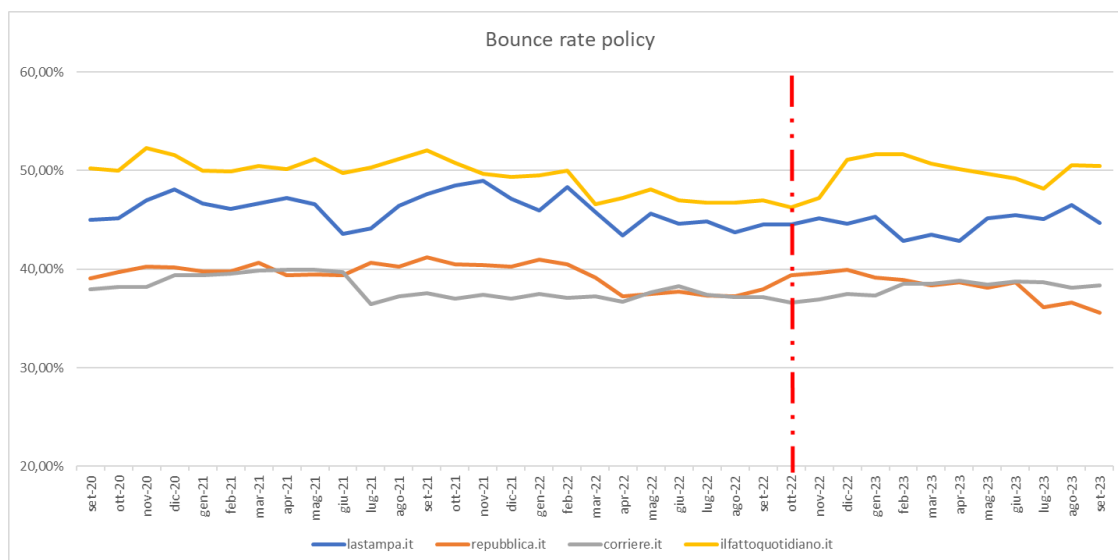


Fig. A.9, Serie temporali Bounce rate giornali policy. Fonte: Elaborazione personale

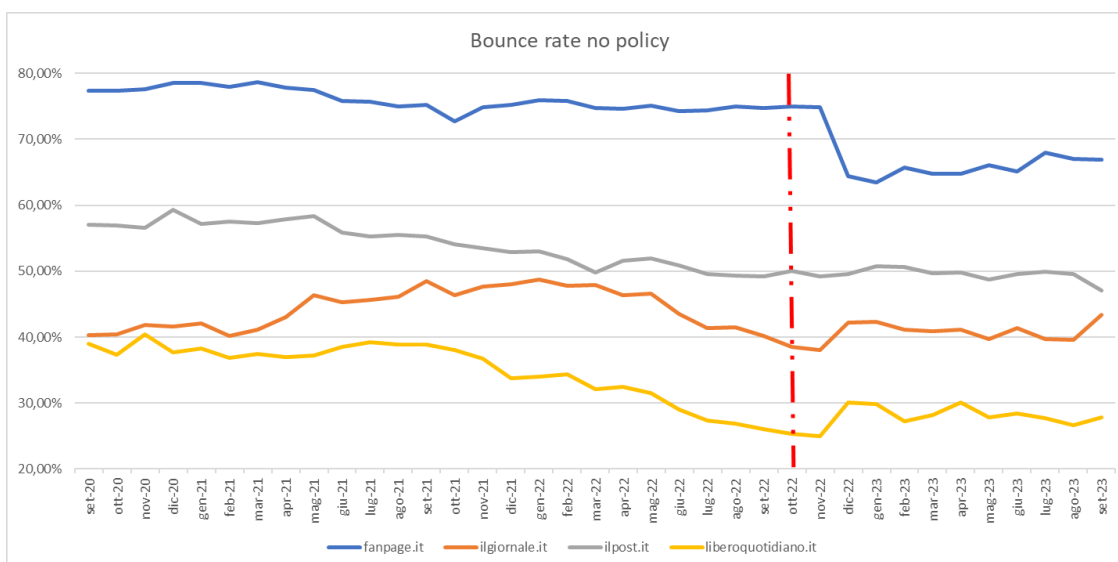


Fig. A.10, Serie temporali Bounce rate giornali no policy. Fonte: Elaborazione personale

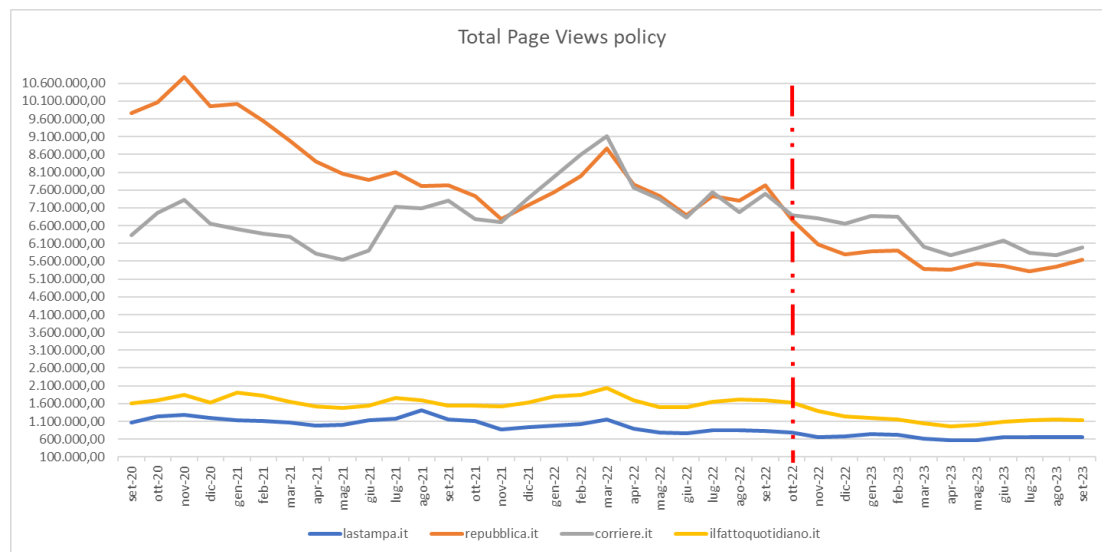


Fig. A.11, Serie temporali Total Page Views giornali policy. Fonte: Elaborazione personale

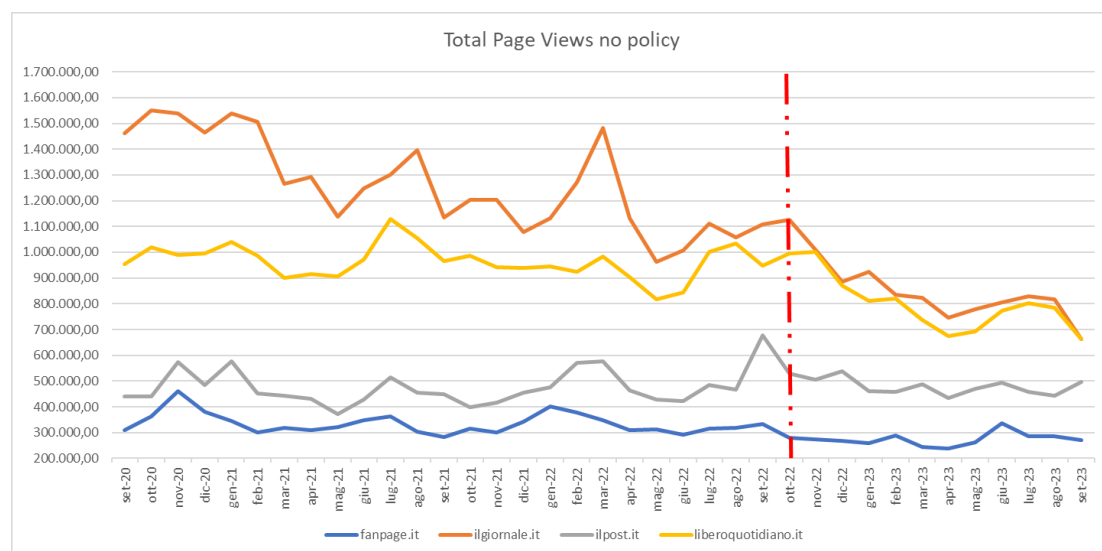


Fig. A.12, Serie temporali Total Page Views giornali no policy. Fonte: Elaborazione personale

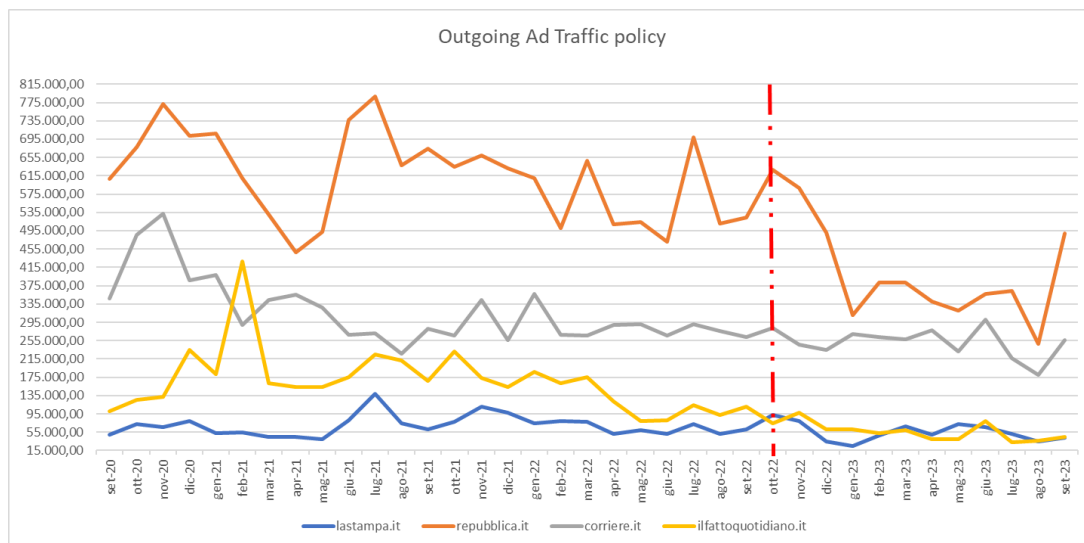


Fig. A.13, Outgoing Ad Traffic giornali policy. Fonte: Elaborazione personale

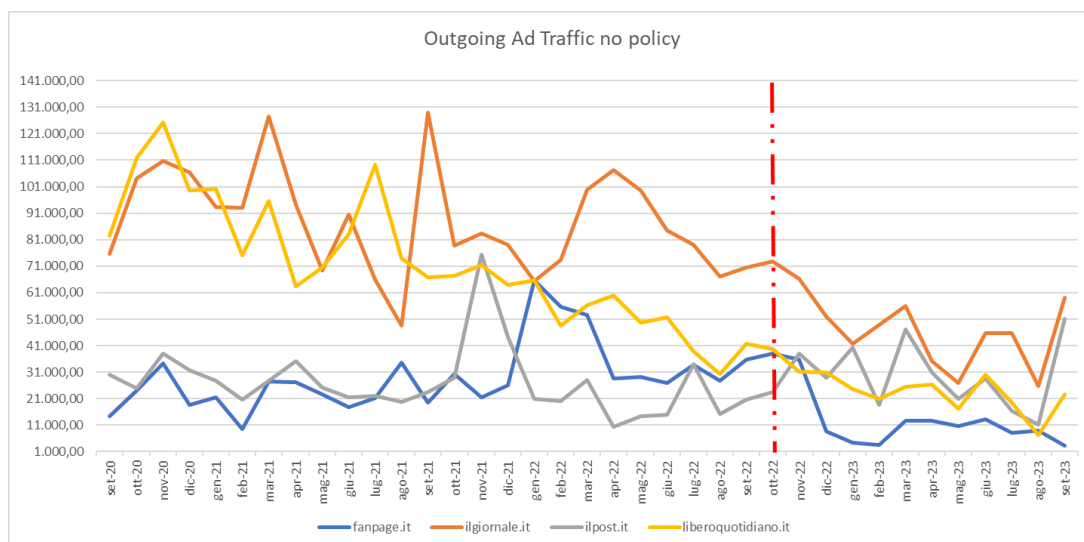


Fig. A.14, Outgoing Ad Traffic giornali no policy. Fonte: Elaborazione personale

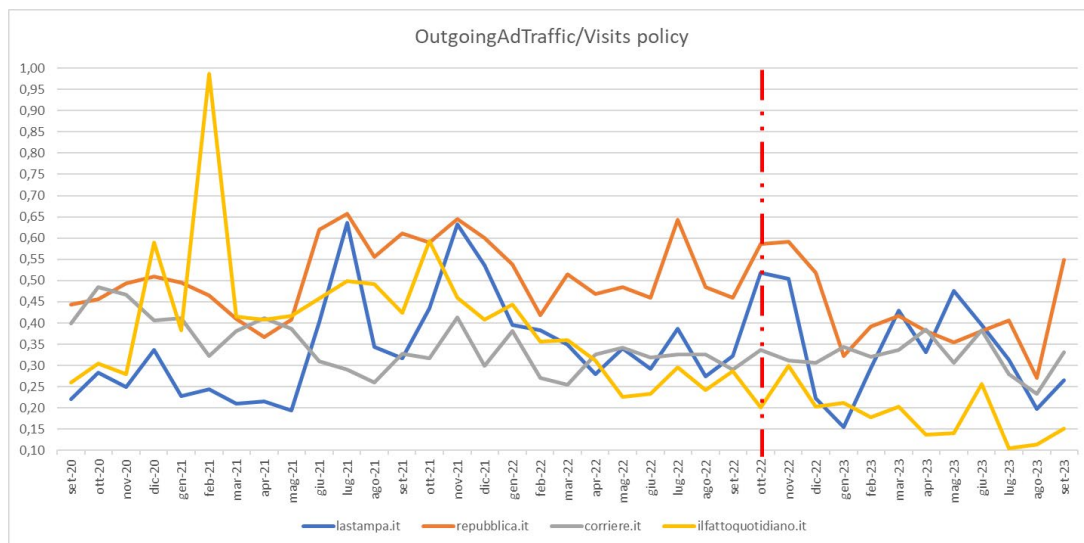


Fig. A.15, OutgoingAdTraffic/Visits giornali policy. Fonte: Elaborazione personale

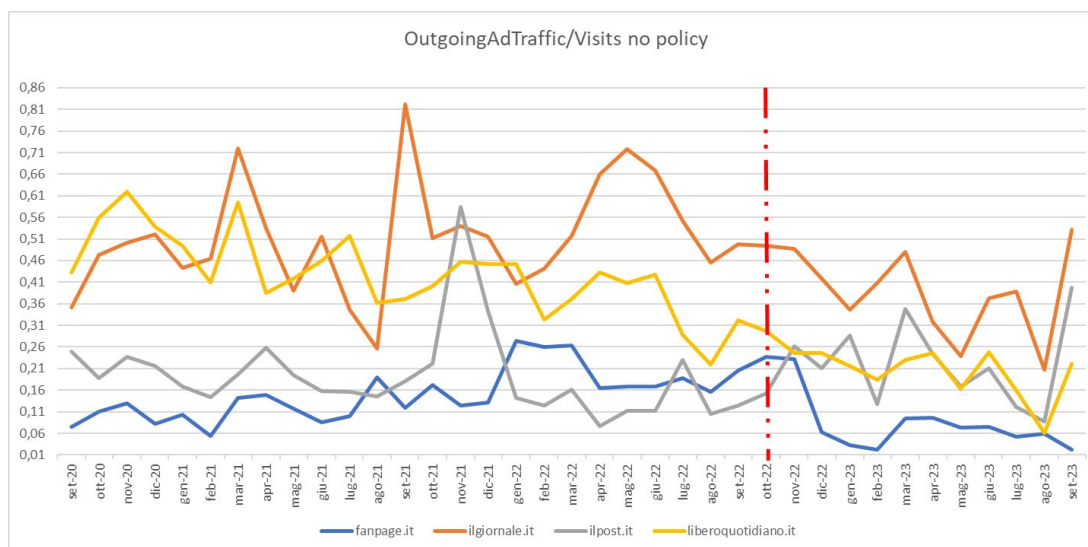


Fig. A.16, OutgoingAdTraffic/Visits giornali no policy. Fonte: Elaborazione personale

BIBLIOGRAFIA

Aridor, G., Che, Y. K., & Salz, T. (2020). The economic consequences of data privacy regulation: Empirical evidence from gdpr (pp. 1-67). Cambridge, MA, USA: National Bureau of Economic Research.

Armstrong, M. (2006). Competition in two-sided markets. *Rand Journal of Economics*, 37(3), 668-691.

AUDICOM, (2023). Sistema Audipress: dati lettura dei quotidiani e dei periodici in Italia 2023 II.

AGCOM (2023). Focus Bilanci 2017-2021.

Bergemann, D., & Bonatti, A. (2015). Selling cookies. *American Economic Journal: Microeconomics*, 7(3), 259-294.

Bettini, A. (2006). Giornali. it: la storia dei siti internet dei principali quotidiani italiani. *Giornali. it*, 0-0.

Bhargava, H. K. (2014). Platform technologies and network goods: insights on product launch and management. *Information Technology and Management*, 15, 199-209.

Brozzetti, E., & Segatori, R (2015). IL GIORNALISMO ON LINE. IL CASO DEL POST.

Caillaud, B., & Jullien, B. (2003). Chicken & egg: Competition among intermediation service providers (Vol. 34, pp. 309-328).

Chemolli, E., & Pasini, M. (2008). I dati mancanti. *DiPAV-QUADERNI*, (2007/20).

COMMISSIONE UE (2015), Public consultation on the regulatory environment for platforms, online intermediaries, data and cloud computing and the collaborative economy.

COMMISSIONE UE (2016), Comunicazione della Commissione al Parlamento europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle Regioni, Le piattaforme online e il mercato unico digitale Opportunità e sfide per l'Europa.

Colone, B. (2020). Big data analytics e gestione aziendale: il nuovo oro? Focus sui cambiamenti apportati nel settore dei trasporti aerei.

Crémer, J., de Montjoye, Y. A., & Schweitzer, H. (2019). Competition policy.

Cutrupi, A. (2017). Piattaforme digitali: tra criticità concorrenziali e prospettive di regolamentazione.

Evans, D. (2003). Some Empirical Aspects of Multi-sided Platform Industries. *Review of Network Economics*, 2(3), 2194-5993.

Evans, D.S., & Schmalensee, R. (2007). The Industrial Organization of Markets with TwoSided Platforms *Competition Policy International*, 3(1), 151-179.

Evans, D. S., & Schmalensee, R. (2010). Failure to launch: Critical mass in platform businesses. *Review of network economics*, 9(4).

Evans, D. S., & Schmalensee, R. (2019). Antitrust analysis of platform markets: Why the supreme court got it right in American express. *Antitrust Analysis of Platform Markets: Why the Supreme Court Got It Right in American Express*, *Competition Policy International*.

Evans, P. C., & Gawer, A. (2016). The rise of the platform enterprise: A global survey.

Ferrazza, F. (2016) 9 ragioni per cui Corriere.It a pagamento ce la può fare. Oppure no.

Filistrucchi, L., Geradin, D., & Van Damme, E. (2013). Identifying Two-Sided Markets. *World Competition: Law & Economics Review*, 36(1), 33-60.

Franceschini, E. (2009) La svolta firmata Murdoch 'I nostri siti presto a pagamento'.

Fiore, C. (2023). Paywall e cookie wall: l'Autorità Garante tedesca prende posizione. *Diritto al digitale*.

Gawer, A. (2014). Bridging differing perspectives on technological platforms: Toward an integrative framework. *Research policy*, 43(7), 1239-1249.

Giacheri Fossati, L. (2005). *Enciclopedia dei ragazzi*.

Goddard, M. (2017). The EU General Data Protection Regulation (GDPR): European regulation that has a global impact. *International Journal of Market Research*, 59(6), 703-705.

Goldfarb, A., & Tucker, C. (2019). Digital economics. *Journal of economic literature*, 57(1), 3-43.

He, W., Li, M., & Zheng, J. (2023). Switching cost, network externality and platform competition. *International Review of Economics & Finance*, 84, 428-443.

Humby, C. (2006). Data is the new oil. *Proc. ANA Sr. Marketer's Summit*. Evanston, IL, USA.

Iadecola, L. (2022). I cookie wall sotto l'occhio del Garante privacy: pratica legittima? *Il Quotidiano Giuridico*.

Laney, D. (2001). 3D Data Management: Controlling Data Volume, Velocity, and Variety. Application Delivery Strategies.

Magrini, D. (2002). Sbatti il web in prima pagina: dati e opinioni. Milano: Franco Angeli.

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Hung Byers, A. (2011). Big data: The next frontier for innovation, competition, and productivity. McKinsey Global Institute.

Martens, B. (2016). An economic policy perspective on online platforms. Institute for Prospective Technological Studies Digital Economy Working Paper, 5.

Mayer-Schönberger, V., & Cukier, K. (2013). Big data: una rivoluzione che trasformerà il nostro modo di vivere e già minaccia la nostra libertà. Garzanti.

Muhtaroglu, F. C. P., Demir, S., Obalı, M., & Girgin, C. (2013, October). Business model canvas perspective on big data applications. In 2013 IEEE International Conference on Big Data (pp. 32-37). IEEE.

Myllylahti, M. (2020). Paying attention to attention: A conceptual framework for studying news reader revenue models related to platforms. *Digital Journalism*, 8(5), 567-575.

Nachbar, T. B. (2021). Platform Effects. *JURIMETRICS J.*, 62, 1.

Regolamento UE 2016/679 (GDPR – /General Data Protection Regulation/) e D.Lgs. 30/6/2003, n. 196 (Codice in materia di protezione dei dati personali).

Richeri, G. (2012). *Economia dei media*. Gius. Laterza & Figli Spa.

Rochet, J.-C., & Tirole, J. (2003). Platform Competition in Two-sided Markets. *Journal of the European Economic Association*, 1(4), 990-1029.

Rochet, J.-C., & Tirole, J. (2006). Two-sided markets: a progress report. *Rand Journal of Economics*, 37(3), 645-667.

Rußell, R., Berger, B., Stich, L., Hess, T., & Spann, M. (2020). Monetizing online content: Digital paywall design and configuration. *Business & Information Systems Engineering*, 62, 253-260.

Rysman, M. (2009). The economics of two-sided markets. *Journal of economic perspectives*, 23(3), 125-143.

Sabatino, L. (2023). Economics of platforms. Slide del corso di Economia e Management dei Servizi. Politecnico di Torino.

Siapera, E. (2019). I dilemmi del giornalismo: le sfide di internet per il giornalismo professionale e la sostenibilità dei media. In *Il giornalismo sotto attacco* (pp. 303-352). Nomos Verlagsgesellschaft mbH & Co. KG.

Spulber, D. F. (2010). Solving the circular conundrum: communication and coordination in internet markets. *Nw. UL Rev.*, 104, 537.

Veisdal, J. (2020). The dynamics of entry for digital platforms in two-sided markets: a multi-case study. *Electronic markets*, 30(3), 539-556.

SITOGRAFIA

www.audiweb.it

www.avangarde-software.com

www.extraordy.com

www.ilmessaggero.it

www.larepubblica.it

www.wired.it