AUTORITÀ GARANTE DELLA CONCORRENZA E DEL MERCATO

AUTORITÀ PER LE GARANZIE NELLE COMUNICAZIONI

GARANTE PER LA PROTEZIONE DEI DATI PERSONALI

INDAGINE CONOSCITIVA SUI *BIG DATA*

Sommario

Premessa

Il 30 maggio 2017 l'Autorità per le garanzie nelle comunicazioni (di seguito "AGCOM"), con delibera n. 217/17/CONS recante "*Avvio di un'indagine conoscitiva sui big data*"[1], l'Autorità garante della concorrenza e del mercato (di seguito "AGCM"), con provvedimento n. 26620 del 30 maggio 2017 "*IC53 – Big Data*", e il Garante per la protezione dei dati personali (di seguito "Garante") – sulla base delle determinazioni adottate nell'adunanza collegiale dell'11 maggio 2017, hanno avviato congiuntamente una Indagine conoscitiva volta ad approfondire la conoscenza degli effetti prodotti dal fenomeno dei *Big Data* e analizzarne le conseguenze in relazione all'attuale contesto economico-politico-sociale e al quadro di regole in vigore.

Nell'ambito dell'Indagine, l'AGCOM ha presentato in un *Interim Report*[2], pubblicato a giugno 2018 sul proprio sito web, le prime considerazioni sulle caratteristiche e sull'ecosistema dei *Big Data*, nonché sul valore economico dei dati e su come le *app* gestiscono i dati e i permessi di accesso. Nel documento, in particolare, si è evidenziato come "*I fallimenti di mercato si ripercuotono su tutto il contesto sociale, compreso il sistema dell'informazione, il pluralismo delle fonti, e le stesse modalità di aggregazione sociale e di formazione dell'opinione pubblica. In conseguenza dell'esistenza di strutturali e duraturi fallimenti di mercato, è necessario, soprattutto laddove sono in discussione diritti sociali e politici, adottare un approccio ex ante alla regolamentazione del dato (e ai connessi algoritmi). Peraltro, questo nuovo paradigma deve considerare che le asimmetrie informative tra utenti e operatori sono pervasive e strutturali. In questo contesto, è difficile ripristinare condizioni di efficienza attraverso meccanismi di trasparenza e di consenso informato. Infatti, tali strumenti appaiono, in molti casi, insufficienti a garantire un riequilibrio conoscitivo tra operatori e consumatori, in una situazione in cui spesso soggetti quali esperti del settore, istituzioni specializzate, nonché centri di ricerca non hanno a disposizione elementi conoscitivi sufficienti a comprendere l'entità e la natura stessa dei fenomeni. In linea con quanto avviene già in contesti ad alta tecnologia (quali quelli delle comunicazioni elettroniche), appare necessario accompagnare la nuova regolazione verso forme tecniche di regolazione diretta degli operatori che utilizzano i Big Data*."[3].

Contemporaneamente, l'AGCM ha pubblicato sul proprio sito istituzionale, nel mese di giugno 2018, i risultati dell'"*Analisi della propensione degli utenti online a consentire l'uso dei propri dati a fronte dell'erogazione di servizi*"[4]. In particolare, il sondaggio, condotto su un campione di utenti di servizi *online*, ha esaminato tre temi: i) il grado di consapevolezza degli utenti delle piattaforme digitali in relazione alla cessione e all'utilizzo dei propri dati individuali; ii) la disponibilità degli utenti a cedere i propri dati personali come forma di pagamento dei servizi *online*; iii) la portabilità dei dati da una piattaforma all'altra. In sintesi, è stato rilevato che circa 6 utenti su 10 non solo sono consapevoli di generare, con le loro attività *online*, dati utilizzabili per attività di profilazione, ma anche che essi

[1] L'avvio di tale indagine è successivo all'indagine conoscitiva avviata dall'AGCOM con la delibera n. 357/15/CONS, riguardante lo sviluppo delle piattaforme digitali e dei servizi di comunicazione elettronica. Una prima parte dell'indagine conoscitiva, dedicata ai cosiddetti "*consumer communications services* (le applicazioni che consentono lo scambio di contenuti vocali, messaggi, foto e video fra due o più utenti), si è conclusa con l'approvazione della delibera n. 165/16/CONS recante "*Indagine conoscitiva concernente lo sviluppo delle piattaforme digitali e dei servizi di comunicazione elettronica di cui alla delibera n. 357/15/CONS: proroga dei termini e pubblicazione della parte relativa ai "consumer communications services*". Una seconda parte, concernente lo sviluppo delle piattaforme digitali, è confluita nella presente indagine.

[2] "Big data - Interim report" nell'ambito dell'indagine conoscitiva di cui alla delibera n. 217/17/CONS.

[3] *Ibidem* pag. 7.

[4] https://www.agcm.it/dotcmsDOC/allegati-news/IC53%20-%20Survey%20primi%20risultati.pdf

appaiono informati dell'elevato grado di pervasività dei sistemi di raccolta (es. geo-localizzazione, accesso a funzionalità come la rubrica, il microfono e la videocamera) e della possibilità di sfruttamento dei dati da parte delle imprese. Nel complesso è risultato che 4 utenti su 10 sono consapevoli della stretta relazione esistente tra la concessione del consenso e la gratuità del servizio. Dal sondaggio è emerso altresì che solo 1 utente su 10 è consapevole dei propri diritti in materia di portabilità dei dati e che circa la metà degli utenti mostra interesse ad ottenere una copia dei propri dati. Il basso interesse all'utilizzo della portabilità è dovuto alla scarsa propensione ad utilizzare altre piattaforme/applicazioni (41,1%), ad una limitata sensibilità sulla rilevanza di tali dati (36,1%), nonché alla percezione di un'elevata complessità degli strumenti tecnologici (30,4%).

Nel presente documento, in continuità con i suddetti *paper*, vengono riportati i risultati dell'indagine derivanti dall'elaborazione delle risposte alle richieste di informazioni e dei contributi forniti dai numerosi operatori[5] ed esperti[6] sentiti in audizione, nell'arco temporale compreso tra novembre 2017 e novembre 2018, nonché dall'acquisizione di conoscenze provenienti da altre fonti, quali la letteratura in materia (economica e giuridica, anzitutto) e testi normativi.

La presente Indagine conoscitiva è articolata in 5 capitoli e un capitolo conclusivo. Il capitolo 1 introduce i temi oggetto dell'Indagine e fornisce una definizione e una descrizione delle caratteristiche dei *Big Data*. Nel capitolo 2 vengono riportate le principali questioni emerse nel corso delle audizioni e dai contributi dei partecipanti all'Indagine e i riflessi sull'operatività delle imprese italiane. Il capitolo 3 riporta le considerazioni dell'AGCOM su come il fenomeno dei *Big Data* incida nel settore delle comunicazioni elettroniche e dei media. Il capitolo 4 riporta le considerazioni del Garante sul possibile impatto dei *Big Data* sul diritto alla protezione dei dati personali e sulle misure e cautele da adottare; il capitolo 5 quelle dell'AGCM sull'utilizzo dei *Big Data* e le relative implicazioni di natura *antitrust* e di tutela del consumatore. Infine, nel capitolo conclusivo sono descritte le linee guida e raccomandazioni di *policy*.

1. *I Big Data*

1.1. Introduzione ai *Big Data*

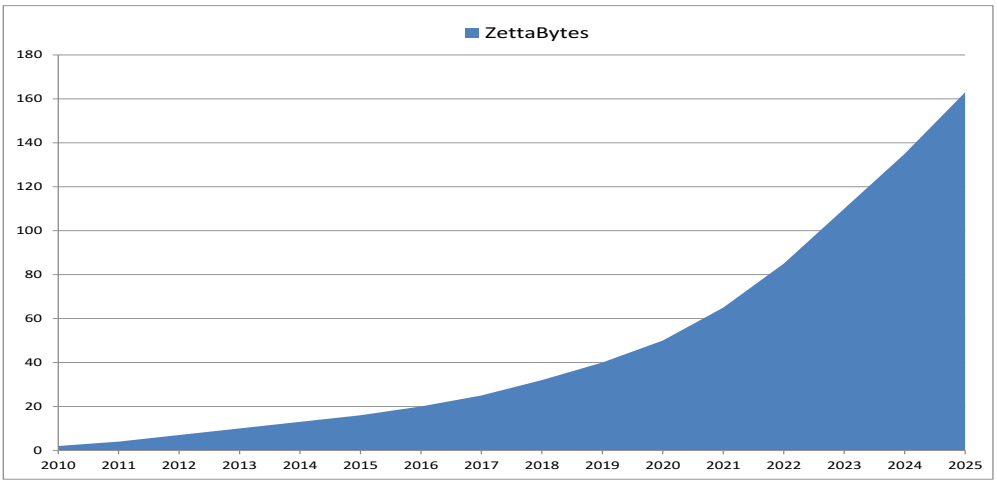
Negli ultimi anni i dati hanno assunto importanza via via crescente nell'organizzazione delle attività di produzione e di scambio, a tal punto da poter essere considerati una risorsa economica a tutti gli effetti, anzi la risorsa di gran lunga più importante in molti settori. Infatti, grazie agli avanzamenti nell'ambito dell'*Information e Communication Technology* (ICT), le organizzazioni tendono a raccogliere dati di qualsiasi tipo, ad elaborarli in tempo reale per migliorare i propri processi decisionali e a memorizzarli in maniera permanente al fine di poterli riutilizzare in futuro o di estrarne nuova conoscenza.

La creazione di dati sta seguendo un processo esponenziale: nell'anno 2018 il volume totale di dati creati nel mondo è stato di 28 zettabyte (ZB)[7], registrando un aumento di più di dieci volte rispetto al

[5] Si sono tenute 18 audizioni con soggetti che operano nei seguenti settori: Telecomunicazioni (Tim, Vodafone, Fastweb Wind Tre), Media (RAI, Mediaset, GEDI, Sole24ore), OTT (Facebook, Microsoft, Amazon), Information Technology (CRIF, Experian, IBM), Assicurazioni e Credito (Allianz, Intesa San Paolo, Unicredit, Generali).
[6] Si sono tenute 8 audizioni con professori universitari ed esperti del settore: Politecnico di Torino, Università degli Studi di Milano, Università Commerciale Luigi Bocconi, Università di Pisa (KDD Lab), Università La Sapienza di Roma, Università Europea di Roma, Università degli Studi di Napoli Parthenope, Université de Namur, University of Malta.
[7] Con 1 ZB pari a un trilione di gigabyte, ovvero 250.000.000.000 di DVD.

2011[8]; come rappresentato in Figura 1, si prevede che entro il 2025 il volume complessivo dei dati arriverà fino a 163 ZB.

Figura 1 – La crescita dei dati

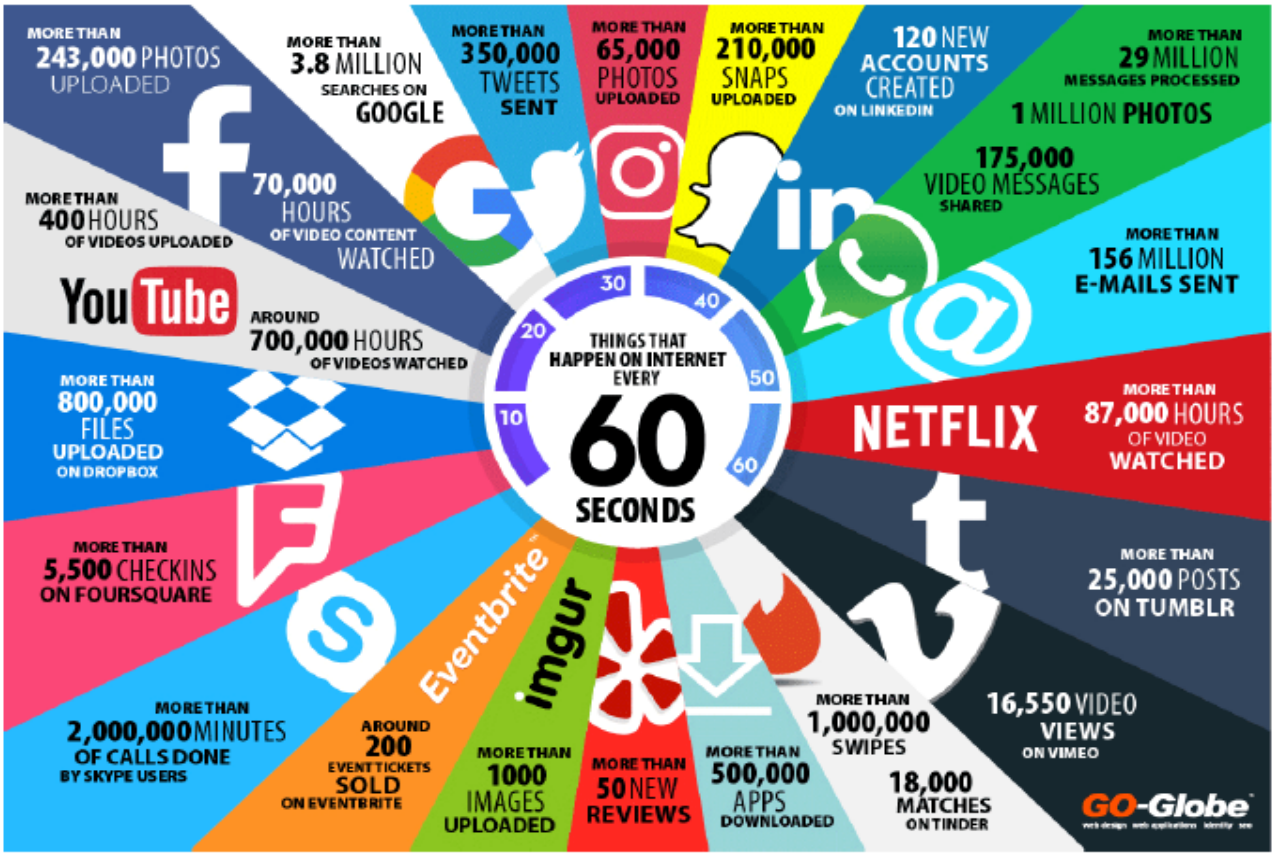


Fonte: elaborazione AGCM in base ai dati forniti nel rapporto tecnico IDC[8]

Il motore primario di questo processo di generazione di dati è indubbiamente Internet: attraverso la rete, infatti, in un minuto sono inviati 44 milioni di messaggi, sono effettuate 2,3 milioni di ricerche su Google, sono generati 3 milioni di "mi piace" e 3 milioni di condivisioni su Facebook, e sono effettuati 2,7 milioni di *download* da YouTube. Google elabora dati di centinaia di Petabyte (PB), Facebook ne genera oltre 10 PB al mese e Alibaba decine di Terabyte (TB) al giorno per il commercio *online* (cfr. Figura 2).

[8] Cfr. Rapporto tecnico dell'International Data Corporation (IDC): David Reinsel, John Gantz, John Rydning. "Data Age 2025: *The Evolution of Data to Life-Critical. Don't Focus on Big Data; Focus on the Data That's Big*". *IDC Report*, 2017.

Figura 2 – Le informazioni generate su Internet in un minuto



Fonte: Go-Globe.com[9]

1.2. Definizioni

In tale contesto, con la locuzione "*Big Data*" si fa riferimento, in prima approssimazione (nell'assenza di definizioni normativamente vincolanti), alla raccolta, all'analisi e all'accumulo di ingenti quantità di dati, tra i quali possono essere ricompresi dati di natura personale (nell'accezione fornita dall'art. 4 del Regolamento (UE) 2016/679 del Parlamento europeo e del Consiglio del 27 aprile 2016, relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati e che abroga la direttiva 95/46/CE, di seguito anche "RGPD"), in ipotesi provenienti anche da fonti diverse. La natura massiva delle operazioni di trattamento reca con sé la necessità che tali insiemi di informazioni (sia memorizzate, sia in streaming) siano oggetto trattamento automatizzato, mediante algoritmi e altre tecniche avanzate, al fine di individuare correlazioni di natura (per lo più) probabilistica, tendenze e/o modelli[10].

Operativamente, nel settore dell'ICT, per *Big Data* si intende una collezione di dati che non può essere acquisita, gestita ed elaborata da strumenti informatici, da *software* e da *hardware* "tradizionali" in un tempo tollerabile[11], benché non esista una soglia dimensionale predefinita affinché un insieme di dati possa essere ricondotto alla categoria dei *Big Data* (per esempio anche un

[9] https://www.go-globe.com/blog/things-that-happen-every-60-seconds/

[10] Così la Risoluzione del Parlamento Europeo del 14 marzo 2017 sulle implicazioni dei *Big Data* per i diritti fondamentali: privacy, protezione dei dati, non discriminazione, sicurezza e attività di contrasto (2016/2225(INI)). Per una prima ricognizione che ha messo in luce le potenzialità (oltre che la rilevanza economica), ma anche i rischi dei Big Data, si rinvia a D. Bollier, The Promise and Peril of Big Data, Washington, DC, 2010.

[11] The McKinsey Global Institute, 2012.

insieme di dati di qualche GB, in cui ogni record è composto da più di 500 mila variabili, può essere definito *Big Data*, se algoritmi tradizionali di analisi non riescono a computare un risultato in tempi ragionevoli su un computer di fascia alta).

In chiave descrittiva è frequente rinvenire nella letteratura in materia[12], fortemente influenzata dall'esperienza nord-americana, il richiamo, in forma ellittica, ad alcune caratteristiche ricorrenti rispetto al fenomeno in esame. Esse sono sintetizzate nelle 4 "V": il **volume**, con riferimento all'enorme dimensione dei dati generati e raccolti; la **varietà**, con riguardo alle numerose tipologie dei dati disponibili, tra i quali, oltre ai dati strutturati tradizionali, vi sono anche dati semi-strutturati e non strutturati come audio, video, pagine *web* e testi; la **velocità** delle operazioni di trattamento; il **valore** che i dati assumono allorquando vengono elaborati ed analizzati, così da consentire l'estrazione di informazioni che possono contribuire all'efficienza e alla qualità di processi produttivi "tradizionali" ovvero qualificare intrinsecamente l'offerta di beni e/o servizi, in particolare in termini di innovazione e di personalizzazione[13].

Sono state poi individuate molteplici altre V idonee a caratterizzare i *Big Data*: tra le più degne di nota si ricordano la *veridicità*, ovvero la qualità e significatività dei dati raccolti o elaborati, la *valenza,* ovvero il grado di connessione del dato con altri dati, la *visualizzazione*, ovvero la necessità di riassumere in maniera visuale e facilmente interpretabile i dati più rilevanti e la conoscenza estratta da essi.

1.3. La filiera dei *Big Data*

Finalità ultima degli articolati processi sottesi all'utilizzo di *Big Data* vuole essere, in termini generali, quella di "accrescere l'efficienza dei processi produttivi, migliorare la capacità decisionale degli amministratori, prevedere più accuratamente le tendenze di mercato e indirizzare in modo molto più mirato (e dunque variamente efficiente) la pubblicità o le diverse proposte commerciali"[14]

A tal fine è cruciale il processo di "estrazione di conoscenza" dai *Big Data*, nell'ambito del quale è possibile enucleare, sul piano logico (con possibili ricadute anche su quello giuridico) tre ordini principali di attività: i) la raccolta, che a sua volta si articola in generazione, acquisizione e memorizzazione, ii) l'elaborazione, che coinvolge attività di estrazione, integrazione e analisi, iii) l'interpretazione e l'utilizzo (Cfr. Figura 3). Ciascuna di esse sarà approfondita nei successivi paragrafi.

---

[12]D. Laney. "3D Data Management: controlling data Volume, Velocity and Variety", *META Group Report*, File 949, 2001. Successivamente, nel 2012, in un nuovo report venne coniata la definizione. M.A. Beyer e D. Laney. "The importance of Big data: a Definition", *Gartner Analysis Report*, ID: G00235055, 2012 e M.A. Beyer. "Gartner says solving big data challenge involves more than just managing volumes of data." *Gartner Report*, 2011, http://www.gartner.com/it/page.jsp. Sul punto si vedano anche J. Gantz, D. Reinsel, "*Extracting value from chaos*" *IDC Report*, 2011, Min Chen, Shiwen Mao, and Yunhao Liu, "*Big data: A survey*" *Mobile networks and applications*, 19.2: 171-209, 2014 e OECD - Organization for Economic Co-operation and Development, "*Big data: Bringing competition policy to the digital era*", 2016 http://www.oecd.org/competition/big-data-bringing-competition-policy-to-the-digital-era.htm. Da ultimo, per una rassegna della letteratura, si veda anche M. Delmastro e A. Nicita, *Big data. Come stanno cambiando il nostro mondo*, il Mulino, Bologna 2019.

[13] La rilevanza delle analisi ed elaborazioni ai fini della valorizzazione dei dati è emersa nell'ambito di diverse audizioni svolte nella presente indagine. Sul punto cfr. paragrafo 1.4.

[14] Tali considerazioni si rinvengono già nel citato rapporto interlocutorio realizzato da parte di AGCOM, Big Data. Interim report nell'ambito dell'indagine conoscitiva di cui alla delibera n. 217/17/CONS, giugno 2018. Cfr. anche l'audizione C. Giustozzi del 16 novembre 2017.

Figura 3 - La filiera dei *Big Data*

1. **Raccolta**
   - Generazione
   - Acquisizione
   - Memorizzazione
2. **Elaborazione**
   - Estrazione
   - Integrazione
   - Analisi
3. **Interpretazione**
   - Interpretazione
   - Decisione

Fonte: elaborazione AGCM

Prima di considerare in modo più approfondito tali fasi, è necessario evidenziare che i dati oggetto di elaborazione secondo le tecniche proprie dei *Big Data* possono avere natura personale o non personale[15], distinzione che rileva ai fini del trattamento dei dati sotto il profilo regolamentare. Nell'ipotesi, più frequente, in cui formino oggetto di elaborazione dati non aventi carattere personale (quali, ad esempio le informazioni di natura geografica, meteorologica, ambientale, economica, etc.) trova applicazione il recente Regolamento (UE) 2018/1807 del 14 novembre 2018 relativo a un quadro applicabile alla libera circolazione dei dati non personali nell'Unione europea[16]. D'altra parte, con riguardo al trattamento dei dati di natura personale è stato previsto uno specifico regime di protezione nell'ambito del quadro normativo recentemente definito a livello europeo, a cui concorrono sia il RGPD sia regole speciali per le attività *online*, individuate nella direttiva 2002/58/CE, relativa al trattamento dei dati personali e alla tutela della vita privata nel settore delle comunicazioni elettroniche e, quindi, nella direttiva 2009/136/CE[17].

[15] Salvo quanto si evidenzierà nel testo, in prima approssimazione per dati personali si intendono le informazioni che identificano o rendono identificabile, direttamente o indirettamente, una persona fisica e che possono fornire informazioni su di essa. I dati non personali, invece, rappresentano le informazioni che non sono relative a persone identificate o identificabili.

[16] Secondo il Considerando n. 9 del citato Regolamento "L'espansione dell'Internet degli oggetti, l'intelligenza artificiale e l'apprendimento automatico rappresentano fonti importanti di dati non personali, ad esempio a seguito del loro utilizzo in processi automatizzati di produzione industriale. Fra gli esempi specifici di dati non personali figurano gli insiemi di dati aggregati e anonimizzati usati per l'analisi dei megadati, i dati sull'agricoltura di precisione che possono contribuire a monitorare e ottimizzare l'uso di pesticidi e acqua, o i dati sulle esigenze di manutenzione delle macchine industriali. Se i progressi tecnologici consentono di trasformare dati anonimizzati in dati personali, tali dati sono trattati come dati personali e si applica di conseguenza il Regolamento (UE) 2016/679".

[17] DIRETTIVA 2009/136/CE DEL PARLAMENTO EUROPEO E DEL CONSIGLIO del 25 novembre 2009 recante modifica della direttiva 2002/22/CE relativa al servizio universale e ai diritti degli utenti in materia di reti e di servizi di

In questa prospettiva è necessario che chi intenda effettuare operazioni di trattamento secondo la metodologia propria dei *Big Data* si accerti, in via preliminare, della natura personale o meno dei dati trattati, così da identificare la cornice normativa di riferimento all'interno della quale opera. In questa prospettiva, sebbene la linea di demarcazione tra dati di natura personale e non possa essere in concreto difficile da tracciare, in particolare in ragione della possibilità di riconnettere informazioni apparentemente anonime (o anonimizzate) a individui singoli a seguito delle peculiari operazioni di trattamento effettuate (nel tempo sempre più agevolmente realizzabili, sia per le aumentate capacità di calcolo, sia per la pluralità di archivi in ipotesi utilizzabili, aventi anche genesi ed utilizzi prospettici diversi al tempo della raccolta), un utile contributo può essere ritratto dalle decisioni delle autorità di protezione dei dati – e nell'esperienza italiana, anzitutto, del Garante – e dagli indirizzi assunti dal Comitato europeo[18], come pure dalle migliori prassi via via elaborate (e comunque soggette a continui aggiornamenti) in tema di anonimizzazione dei dati personali[19].

1.3.1. La raccolta dei *Big Data*

La generazione dei dati

Come detto, dal punto di vista descrittivo (ed impregiudicati i vincoli che potrebbero derivare dall'applicazione delle discipline di volta in volta rilevanti), la fase di raccolta dei *Big Data* ha inizio con la *generazione* che si realizza nell'ambito di attività svolte dagli utenti in un contesto informatizzato ovvero nell'ambito della cosiddetta *Internet of things*. Nell'attuale contesto, in cui pressoché tutti i contenuti *media* sono resi disponibili in formato digitale e gran parte delle attività economiche e sociali sono migrate su *internet*, le **attività degli utenti**, sia di tipo *online* che *offline*, possono generare grandi quantità di dati.

In primo luogo, i servizi *online*, spesso popolati dai contenuti degli utenti stessi, costituiscono una grande fonte per i *Big Data* [20]: si pensi, ad esempio, alla posta elettronica, alla navigazione satellitare, ai *social networks*, in cui i fruitori caricano i propri contenuti (foto, video, testi), condividendoli pubblicamente sulle piattaforme digitali, sulle app e sui siti *internet*. A ciò si aggiunge la raccolta dei dati generati dalle funzionalità dei dispositivi personali degli utenti (quali *smartphone*, *tablet* e *personal computer*).

A titolo esemplificativo, la raccolta dati che si realizza nell'ambito dell'attività *online* di un consumatore che effettua un acquisto su un sito di *e-commerce* è stilizzata nella successiva figura 4.

---

comunicazione elettronica, della direttiva 2002/58/CE relativa al trattamento dei dati personali e alla tutela della vita privata nel settore delle comunicazioni elettroniche e del Regolamento (CE) n. 2006/2004 sulla cooperazione tra le autorità nazionali responsabili dell'esecuzione della normativa a tutela dei consumatori. Quest'ultima è in corso di revisione con la proposta di un nuovo regolamento sul rispetto della vita privata e sulla protezione dei dati personali nelle comunicazioni elettroniche: Cfr. Proposta di Regolamento del Parlamento Europeo e del Consiglio relativo al rispetto della vita privata e alla tutela dei dati personali nelle comunicazioni elettroniche e che abroga la direttiva 2002/58/CE (regolamento sulla vita privata e le comunicazioni elettroniche), datata 10 gennaio 2017.

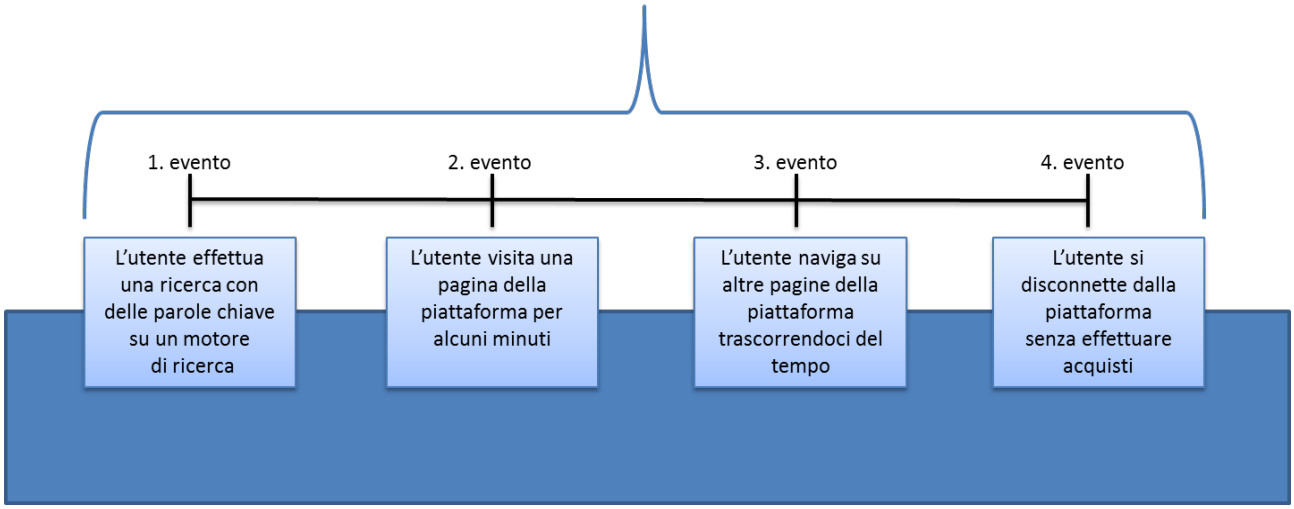
[18] Sulla (ampia) nozione di dato personale si rinvia al WP 136 del Gruppo art. 29, Parere 4/2007 sul concetto di dati personali adottato il 20 giugno 2007. Sul punto cfr. anche, con riguardo al tema dell'anonimizzazione, *infra* par. 4.7.

[19] V. *amplius* le considerazioni svolte al par. 4.

[20] V. già al riguardo, in materia di protezione dei dati personali, i *caveat* contenuti nel WP 148 Parere 1/2008 del Gruppo art. 29 sugli aspetti della protezione dei dati connessi ai motori di ricerca, adottato il 4 aprile 2008, *passim*, ancorché il documento non fosse espressamente dedicato al tema dei *big data*; non diversamente, anche in relazione al trattamento dei dati consentito dalla rilevazione effettuata mediante *cookies*, affinché il trattamento possa considerarsi legittimo deve ora tenersi conto delle condizioni enunciate nella sentenza della Corte di giustizia (Grande Sezione) 1° ottobre 2019, causa C- 673/17, *Bundesverband der Verbraucherzentralen und Verbraucherverbände – Verbraucherzentrale Bundesverband eV c. Planet49 GmbH*.

Figura 4

Il tracciamento delle attività di un consumatore da parte di una piattaforma di e-commerce



Le attività svolte dagli utenti, anche in assenza di interazione diretta con un dispositivo elettronico, generano dati (cosiddetti *offline*) e possono fornire informazioni rilevanti sui comportamenti e sulle preferenze degli individui. Si pensi, ad esempio, ai dati di geolocalizzazione degli individui forniti dagli *smartphone* (nei quali tale funzione risulti attivata), nonostante non vi sia una attiva interazione con il dispositivo da parte dell'utente. Allo stesso modo, le videocamere di sorveglianza, nel riprendere la presenza ed i movimenti degli individui in una determinata zona, acquisiscono dati che poi possono essere elaborati al fine di inferire informazioni sui flussi delle persone. Anche gli strumenti di pagamento elettronici consentono di acquisire informazioni sui comportamenti di acquisto e le preferenze degli utenti che li utilizzano. A questo riguardo vale menzionare l'iniziativa di Google, che ha sottoscritto accordi commerciali con alcuni gestori dei circuiti delle carte di pagamento, al fine di acquisire informazioni sugli acquisiti effettuati dai consumatori, utili a verificare l'efficacia di campagne pubblicitarie personalizzate, nonché a profilare ulteriormente i propri utenti[21].

Un'altra importante fonte per alimentare i *Big Data* è l'internet delle cose (IoT, *Internet of Things*), che vede applicazioni sia in campo industriale (ad esempio nella cosiddetta manutenzione predittiva), sia con riguardo alla vita dei singoli, dalla domotica ai dispositivi, spesso indossabili (*wearable device*) che registrano dati su ogni individuo (ad esempio quelli relativi alle attività sportive e/o ai parametri biologici).

L'idea di base dell'IoT è connettere diversi oggetti del mondo reale - come sensori, attuatori, RFID (Radio-Frequency Identification), lettori di codici a barre, telefoni cellulari, ecc. – e farli cooperare l'uno con l'altro al fine di completare un compito comune, attraverso l'uso di microprocessori presenti negli oggetti[22]. Esso consente lo sviluppo di applicazioni in diversi settori chiave[23]: si pensi, per

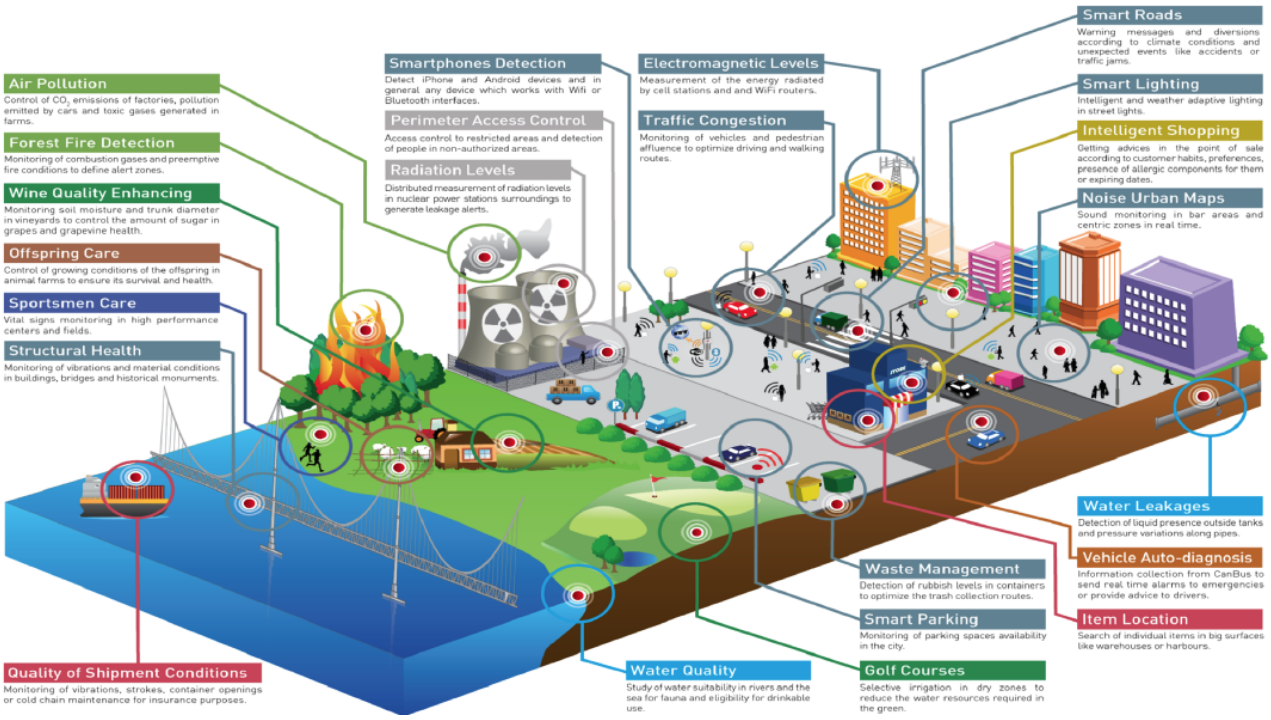
[21] https://www.ilpost.it/2018/08/31/accordo-google-mastercard/;
Douglas C. Schmidt, Professor of Computer Science at Vanderbilt University: "**Google Data Collection**"; DCN 2018. https://digitalcontentnext.org/blog/2018/08/21/google-data-collection-research/

[22] M. Chen, S. Mao, Y. Zhang, V. C. Leung. "*Big data: related technologies, challenges and future prospects*" *Springer*, New York, 2014.

[23] Helen Rebecca Schindler, Jonathan Cave, Neil Robinson, Veronika Horvath, Petal Jean Hackett, Salil Gunashekar, Maarten Botterman, Simon Forge, Hans Graux. "*Europe's policy options for a dynamic and trustworthy development of the Internet of Things*" *RAND, European Union,* 2013.

esempio, ad una Smart City, in cui i cittadini attraverso un'applicazione presente nei propri *smartphone* hanno accesso in tempo reale ai dati sul traffico, sui parcheggi disponibili, sulla qualità dell'aria, sui tempi di attesa dei mezzi pubblici, sulle farmacie di turno aperte, sul numero di pazienti presenti nei pronto soccorsi. Tutto ciò grazie a sensori interconnessi, i quali trasmettono le proprie rilevazioni ad un server centrale che elabora e rende disponibili le informazioni ai propri utenti. La successiva figura mostra uno schema di *smart city*.

Figura 5 - La generazione dei *Big Data*



Fonte: Libelium Smart World - Libelium Comunicaciones Distribuidas S.L[24]

Attraverso i sensori posti nei dispositivi mobili, nei mezzi di trasporto, nelle infrastrutture pubbliche (aeroporti, porti, stazioni ferroviarie) e negli elettrodomestici, l'IoT rende possibile codificare in formato digitale, trasmettere e memorizzare le informazioni afferenti al funzionamento di apparecchiature e dispositivi tra loro connessi, sia in ambito aziendale/lavorativo che con riguardo alle attività dei singoli individui. In tal modo possono essere acquisiti vari tipi di dati (ambientali, geografici e logistici), che, in generale, presentano molteplici caratteristiche tipiche dei *Big Data*, tra cui l'eterogeneità, la varietà, l'assenza di una struttura, la forte relazione spazio/tempo e la rapida crescita.

L'acquisizione dei dati

I dati generati dagli utenti o dalle "cose" nell'ambito dell'IoT vengono quindi acquisiti tramite i dispositivi elettronici coinvolti nell'atto di generazione - quali *smartphones*, sensori (di movimento, di temperatura, di umidità), videocamere, dispositivi di input (tastiera, *mouse*), *scanner*, *RFID*, *wearable devices* e altri dispositivi propri dell'IoT, ecc. – risultando così nella disponibilità dei

FTC Staff Report. "*Internet of things: Privacy & Security in a Connected World*", January 2015.
[24] http://www.libelium.com/libelium-smart-world-infographic-smart-cities-internet-of-things/

soggetti che sviluppano e rendono operativi questi sistemi (ad esempio, nel caso degli *smartphone*, i fornitori del sistema operativo), i quali, tuttavia, per acquisire la disponibilità di dati personali devono necessariamente chiedere il preventivo consenso dell'utente che li ha generati.

Proprio gli *smartphone* rivestono un ruolo centrale nell'acquisizione dei dati generati dagli utenti, in quanto dispongono di numerosi dispositivi di input (come i sensori di movimento, di luminosità, di localizzazione, la tastiera e il *touch screen*) integrati in un unico strumento connesso ad *internet* e che accompagna l'utente in tutte le sue attività quotidiane. Si evidenzia, in particolare, che in uno *smartphone* i veicoli per l'acquisizione dei dati sono rappresentati, da un lato, dal sistema operativo e, dall'altro, dalle applicazioni pre-installate o successivamente acquistate ed installate dall'utente. Nel secondo caso i dati sono acquisiti dai rispettivi sviluppatori.

Più in generale, tutte le attività *online* degli utenti (quali l'invio e la ricezione di posta elettronica, la navigazione satellitare o l'uso di servizi di *social network*) – a prescindere dal dispositivo utilizzato – generano una enorme quantità di dati, tipicamente personali, che alimentano una copiosa attività di acquisizione. A tal proposito, come evidenziato in alcune audizioni svolte nell'ambito della presente Indagine, si consideri il fatto che le modalità di funzionamento dei servizi *online* concorrono a moltiplicare le possibilità di acquisizione di ciascun singolo dato generato con l'utilizzo di *device* elettronici personali[25]. Si sottolinea altresì che, dal lato degli utenti, il "consenso" al trattamento dei dati personali, che le loro attività *online* generano, mira a consentire – ove validamente prestato - la fruizione gratuita delle stesse[26].

Tecnicamente l'acquisizione dei dati generati dagli utenti presuppone l'utilizzo di sistemi dedicati al loro tracciamento che, con specifico riferimento alla navigazione sul *web*, sono costituiti dai cosiddetti *cookie*; questi ultimi sono file di testo che raccolgono le preferenze (es: lingua, interfaccia, luogo dal quale avviene l'accesso, ecc.) e le informazioni del consumatore (es: pagine che ha visitato, testi trasmessi, ecc.) attivo in un sito *web*, consentendone una precisa profilazione, che peraltro viene aggiornata in occasione di ogni successivo accesso al medesimo sito[27].

A tale riguardo si rileva che tra gli sviluppatori di applicazioni e/o di siti *web* è diffusa la prassi di avvalersi *in outsourcing* dei sistemi di tracciamento sviluppati dai principali operatori dell'ICT (quali Apple, Google e Facebook), con la conseguenza che i dati acquisiti dai primi rientrano anche nella disponibilità di questi ultimi[28], che, peraltro, in quanto sviluppatori di sistemi operativi estremamente

[25] Ad esempio, il dottor Quintarelli ha sottolineato che "*ogni volta che [un utente] accede ad una pagina internet, il device utilizzato si collega a molteplici risorse di rete e attiva numerose funzionalità (tecniche e/o di marketing) cosicché diversi soggetti (fornitori di componenti di servizio) acquisiscono dati su quella specifica attività (accesso ad una pagina)*" (13 settembre 2018).

[26] Ma v., in ordine ai requisiti di validità del consenso dell'interessato, da ultimo la sentenza del 29 luglio 2019, Case C- 40/17, *Fashion ID GmbH & Co. KG v Verbraucherzentrale NRW e V*.

[27] In assenza di autenticazione sul sito, i *cookie* costituiscono la più importante modalità di acquisizione dei dati relativi all'utente (oltre al mero tracciamento tramite l'indirizzo IP). I dati in tal modo acquisiti hanno tuttavia un contenuto generalmente più limitato rispetto a quelli che il consumatore rende disponibili quando esegue l'autenticazione al sito.

[28] A questo proposito merita ricordare il caso deciso dalla Corte di giustizia, (Grande Sezione) del 5 giugno 2018, Causa C-210/16 *Unabhängiges Landeszentrum für Datenschutz Schleswig-Holstein c. Wirtschaftsakademie Schleswig-Holstein GmbH*, nel quale è stato affermato che l'articolo 2, lettera d), della direttiva 95/46/CE deve essere interpretato nel senso che la nozione di «titolare del trattamento» (*data controller*), ai sensi di tale disposizione, include (anche) l'amministratore di una *fanpage* presente su un *social network* (oltre al gestore del *social network* medesimo); v. altresì la sentenza del 29 luglio 2019, Case C- 40/17, *Fashion ID GmbH & Co. KG v Verbraucherzentrale NRW e V.*; *Guidelines 2/2019 on the processing of personal data under Article 6(1)(b) GDPR in the context of the provision of online services to data subjects Version 2.0 8 October 2019* e già *Article 29 Working Party Guidelines on consent under Regulation 2016/679* (WP259)

diffusi e/o di app estremamente popolari, sono già in una posizione privilegiata per l'acquisizione diretta dei dati degli utenti dagli *smartphone* e/o dalle relative applicazioni.

Nell'ambito dell'IoT i dati di tipo ambientale, geografico o logistico vengono acquisiti dai dispositivi installati nelle abitazioni ovvero nei siti di produzione industriale, nei locali commerciali e nell'ambiente. Anche per la gestione dei dispositivi dell'IoT è frequente il ricorso a soluzioni standard predisposte dai grandi player ICT (quali, ad esempio, Amazon o Google), che, essendo di regola integrate con funzionalità di elaborazione dei dati acquisiti, assicurano loro la disponibilità dei dati stessi.

Nel processo di acquisizione dei dati possono intervenire anche i cosiddetti *data broker*, ossia soggetti che aggregano dati da diverse fonti (principalmente siti *internet*) e li organizzano per metterli a disposizione di soggetti terzi. Tali intermediari, operando contemporaneamente su molteplici siti, realizzano importanti economie di scala e di scopo (grazie alla varietà dei dati raccolti sui diversi siti) e consentono di aumentare l'ampiezza e la profondità della raccolta dati. I *data broker* alimentano un mercato poco trasparente soprattutto per gli utenti finali, che non sono messi nelle condizioni di conoscere il percorso compiuto dai dati che vengono acquisiti dai siti *internet* e/o dalle piattaforme *online* a cui accedono.

Vi sono, infine, dati che possono essere acquisiti senza doversi interfacciare con gli utenti o comunque con i soggetti che generano quei dati. Si tratta dei cosiddetti *open data,* generalmente prodotti dagli organismi pubblici e per definizione liberamente accessibili a tutti.

La memorizzazione dei dati

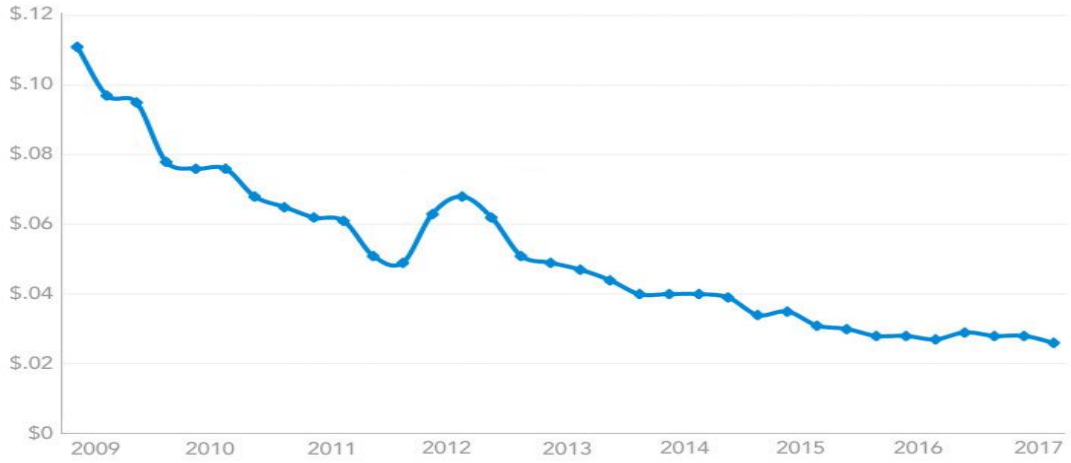
Per *memorizzazione* si intende il processo di trasferimento del dato dal dispositivo di acquisizione alla memoria primaria o secondaria[29] di un sistema di elaborazione in modo tale da poterlo trattare. Nell'ambito di tale processo assume grande rilevanza (specie in ragione delle dimensioni del fenomeno in esame) la dimensione della integrità e sicurezza dei dati.

In considerazione del grande volume dei dati che vengono acquisiti, si rendono necessari, per l'attività di memorizzazione, sistemi di elaborazione dotati di memorie capienti, ad accesso rapido e con tempi di trasferimento veloci. L'accesso a tale risorsa non appare allo stato rappresentare un ostacolo allo sviluppo di attività che coinvolgono i *Big Data*, giacché fino ad ora lo sviluppo tecnologico ha determinato un progressivo trend di riduzione dei prezzi delle memorie (cfr. figura 6).

V. altresì, in ambiti specifici, *Article 29 Working Party Opinion 2/2010 on online behavioural advertising* (WP171); *Article 29 Working Party Working Document 02/2013 providing guidance on obtaining consent for cookies* (WP208).

[29] La memoria secondaria consente una memorizzazione permanente dei dati, diversamente da quella primaria (o volatile), che invece presuppone una elaborazione del dato a brevissima distanza dall'acquisizione.

Figura 6 – Andamento del prezzo per *gigabyte* delle memorie secondarie



Fonte: Backblaze Ltd[30]

1.3.2. L'elaborazione dei *Big Data*

Come riconosciuto dagli esperti sentiti nelle audizioni svolte nell'ambito della presente indagine, i dati isolatamente considerati hanno poco valore, ma lo acquisiscono quando sono organizzati[31]. Per tale ragione riveste un ruolo centrale nell'intera filiera dei *Big Data* la fase della elaborazione, che comporta l'organizzazione dei dati grezzi non strutturati in informazioni suscettibili di essere utilizzate per finalità economiche. L'attività di analisi, infatti, consente di estrarre velocemente conoscenza da grandi moli di dati non strutturati così da ottenere informazioni possibilmente in un formato compatto e facilmente interpretabile.

Dopo una iniziale fase di *estrazione* – durante la quale i dati vengono reperiti dalle diverse fonti disponibili, selezionati e caricati nella memoria del sistema di elaborazione – ed una successiva *integrazione* di tutte le informazioni che si riferiscono agli stessi elementi o domini applicativi, interviene la vera e propria *analisi* dei dati, che avviene per il tramite di tecniche di analisi e strumenti capaci di far emergere dai dati grezzi non strutturati informazioni suscettibili di interpretazione e utilizzo pratico.

In linea generale, le tecniche di analisi consistono per lo più in algoritmi[32] tra i quali si distinguono quelli di interrogazione e quelli di apprendimento. Mentre i primi mirano a rispondere a delle richieste precise da parte degli utenti poste in forma di interrogazioni[33], i secondi invece mirano ad estrarre nuova conoscenza, nuove tesi e si avvalgono di tecniche avanzate di *Intelligenza Artificiale*[34] come il *machine learning*.

---

[30] https://www.backblaze.com/blog/hard-drive-cost-per-gigabyte/
[31] Cfr. ad esempio audizione dei Proff. Preta, Maggiolino e altri (30 gennaio 2018).
[32] Il termine algoritmo indica la sequenza di istruzioni che deve essere effettuata per eseguire un'elaborazione o risolvere un problema.
[33] Ad esempio: restituisci gli acquisti di tutti i consumatori con età minore di 20 e di sesso femminile avvenuti nell'ultimo mese.
[34] La tematica dell'intelligenza artificiale è di crescente attualità ed è oggetto di riflessione anche sui tavoli istituzionali, come emerge peraltro da recenti approfondimenti. Cfr. Comunicazione della Commissione al Parlamento europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle regioni, Creare fiducia nell'intelligenza artificiale antropocentrica, Bruxelles, 8.4.2019, COM(2019) 168 final; Commissione europea per l'efficienza della giustizia (Cepej) del Consiglio d'Europa, *European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their*

La caratteristica di questi algoritmi, il cui funzionamento evolve in base all'esperienza acquisita, è di essere variabili nel tempo, anche con elevata velocità. Inoltre la tendenza ad ottimizzare i modelli computati sulla base dei dati analizzati, li rende sempre più precisi ed accurati. Tali peculiarità rendono gli algoritmi di *machine learning* dotati di una certa "autonomia" di comportamento. Ad esempio, a livello teorico è stato dimostrato come gli *algoritmi di pricing dinamico* basati sull'intelligenza artificiale e sui *Big Data* possono portare a fenomeni di collusione tacita[1], proprio grazie all'apprendimento iterato che ne ottimizza il funzionamento e il modello di *pricing* adottato[35].

Benché generalmente gli algoritmi di analisi dei *Big Data* siano pubblicamente disponibili, quelli effettivamente utilizzati dai singoli operatori finiscono per essere individualizzati e restano sconosciuti ai terzi, se non a grandi linee, giacché ciascun operatore può rilasciare versioni proprietarie attraverso un processo di reingegnerizzazione dei metodi esistenti, così da personalizzare le implementazioni degli algoritmi esistenti e nascondere i dettagli del loro funzionamento agli utilizzatori.

L'implementazione degli algoritmi a sua volta richiede modelli informatici di calcolo che coinvolgono risorse *hardware* e al *software* che nel sempre più diffuso modello del *cloud computing*[36] sono disponibili in *data center remoti* e vengono rilasciate rapidamente e in modo dinamico agli utenti, che le condividono.

Proprio l'"intelligenza" delle tecniche di analisi, unitamente alla voluminosità e varietà dei dati, sta facendo emergere una importante innovazione nel processo di estrazione della conoscenza. Nel nuovo

---

*environmen*t, Strasbourg, 3-4 December 2018 nel quale si sono stabiliti i principi etici relativi all'uso dell'Intelligenza Artificiale (AI), in particolare nei sistemi giudiziari. La Carta intende fornire un quadro di principi destinati a policy maker, legislatori e i professionisti della giustizia con riguardo al rapido sviluppo dell'IA nei procedimenti giudiziari nazionali. L'opinione del CEPEJ, come si evince dalla Carta, è che l'applicazione dell'IA nel campo della giustizia può contribuire a migliorare l'efficienza e la qualità e deve essere attuata in modo responsabile e conforme ai diritti fondamentali garantiti, in particolare nella Convenzione europea sul Diritti umani (CEDU) e la Convenzione del Consiglio d'Europa sulla protezione dei dati personali. Per il CEPEJ, è essenziale garantire che l'IA rimanga uno strumento al servizio dell'interesse generale e che il suo uso rispetti i diritti individuali. In questa prospettiva, il CEPEJ ha identificato i seguenti principi fondamentali da rispettare nel campo dell'IA e della giustizia:

a. principio del rispetto dei diritti fondamentali, al fine di assicurare che la progettazione e l'attuazione di strumenti e servizi di intelligenza artificiale siano compatibili con i diritti fondamentali;
b. principio di non discriminazione, al fine di prevenire lo sviluppo o l'intensificazione di qualsiasi discriminazione tra individui o gruppi di individui;
c. principio di qualità e sicurezza, in relazione al trattamento delle decisioni giudiziarie e dei dati, utilizzando fonti certificate e dati non modificabili con modelli concepiti in modo multidisciplinare, in un ambiente tecnologico sicuro;
d. principio di trasparenza, imparzialità ed equità, al fine di rendere i metodi di trattamento dei dati accessibili e comprensibili, autorizzando audit esterni;
e. principio "under user control" ("sotto il controllo dell'utente"), al fine di prevenire un approccio "prescrittivo" ed assicurare che gli utenti siano attori informati e in controllo delle loro scelte.

Per il CEPEJ, il rispetto di questi principi deve essere assicurato nell'elaborazione delle decisioni giudiziarie e dei dati mediante algoritmi e nell'uso fatto degli stessi.

Ulteriori materiali di approfondimento possono essere rinvenuti al link: https://www.coe.int/en/web/cepej/cepej-european-ethical-charter-on-the-use-of-artificial-intelligence-ai-in-judicial-systems-and-their-environment.

[35] Cfr. A.V. den Boer. "*Dynamic pricing and learning: historical origins, current research, and new directions*." *Surveys in operations research and management science*, 20.1:1-18, 2015. https://doi.org/10.1016/j.sorms.2015.03.001; CMA – Competition & Markets Authority. "*Pricing algorithms. Economic working paper on the use of algorithms to facilitate collusion and personalised pricing*", CMA94, 2018. https://www.gov.uk/government/publications/pricing-algorithms-research-collusion-and-personalised-pricing.

[36] Anche in relazione al *cloud computing*, in caso di trattamenti di dati personali, può essere necessaria l'adozione di adeguate cautele: v. in merito, le indicazioni contenute nel Parere 05/2012 sul *cloud computing* adottato dal gruppo art. 29 il 1° luglio 2012 WP 196, *sub* https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2012/wp196_it.pdf.

paradigma analitico, cosiddetto *data driven*, i dati concorrono non solo a verificare ipotesi teoriche con tecniche statistiche, ma anche a esplorare nuovi scenari e ricavare nuove teorie, nonché, più in generale, a scoprire nuova conoscenza attraverso gli algoritmi di intelligenza artificiale[37]. Si tratta di un approccio all'acquisizione delle informazioni e alla generazione di conoscenza del tutto innovativo dal punto di vista metodologico, che riconosce ai dati il ruolo di guida e agli algoritmi il compito di trovare modelli che la metodologia tradizionale forse solo a fatica potrebbe individuare (salvo doverli poi sottoporre a successiva verifica). La portata innovativa è tale che alcuni studiosi parlano di vera e propria rivoluzione scientifica rispetto all'approccio classico "ipotesi, modello, esperimento"[38].

Nell'ambito di questo nuovo paradigma analitico, i dati appaiono rivestire rilevanza centrale. Infatti – come è stato evidenziato in audizione dai rappresentanti di KDD Lab – "*i programmi di intelligenza artificiale [...] apprendono grazie alla disponibilità di un elevatissimo numero di esempi*". Pertanto, il dato, in quanto sorgente di informazione sul fenomeno che si intende studiare, rappresenta l'origine stessa dell'evoluzione degli algoritmi, cosicché è la disponibilità di nuove fonti di dati che consente il miglioramento degli algoritmi impiegati e/o lo sviluppo di nuovi algoritmi.

D'altra parte, anche quando gli algoritmi non mutano nel tempo, il progresso della conoscenza dipende dai dati. Ad esempio, in diversi ambiti (quali le previsioni metereologiche o la traduzione *online*) i miglioramenti registrati negli ultimi anni sono riconducibili non tanto agli algoritmi, che sostanzialmente non sono mutati rispetto al passato, quanto alla disponibilità di immensi quantitativi di dati, oltre che alla capacità computazionale alquanto più potente[39].

Infine, va considerato che alcune attività tipiche del contesto digitale hanno senso sotto il profilo economico solo se si basano su una grande quantità e varietà di dati (ad esempio, *recommendation system* sulla cd. *long tail*[40] delle piattaforme di vendita *online*).

Al riguardo le società dell'ICT sentite in audizione hanno per un verso sottolineato che la valenza dei dati è inversamente proporzionale alla loro genericità, dal momento che, generalmente, il vantaggio nello sviluppo di soluzioni intelligenti ai problemi di un particolare utente deriva proprio dall'analisi dei dati prodotti al suo interno: "*[..] i dataset più rilevanti per un'impresa sono quelli che l'impresa crea per sé stessa in quanto essa conosce il contesto nel quale sono stati creati e le finalità per le quali erano creati; le principali innovazioni possono derivare proprio dai* dataset *costruiti da un'impresa per uso interno, non destinati fin dall'inizio a terzi o al mercato*" [41]. Per altro verso, gli stessi operatori hanno osservato come "*la precisione degli algoritmi aumenta con la diversità delle fonti di dati cosicché una fonte di dati debolmente correlata ad un fenomeno può avere un impatto maggiore in termini di miglioramento dell'algoritmo di una fonte più precisa e raffinata strettamente connessa al medesimo fenomeno*"[42].

In ogni caso, i colossi dell'economia digitale (quali Google, Apple, Facebook, Amazon, Microsoft) appaiono godere di un vantaggio rispetto alle imprese dei settori tradizionali dal momento che, oltre

[37] Particolari algoritmi che auto-apprendono e migliorano il proprio funzionamento sulla base dell'esperienza ottenuta in iterazioni successive

[38] S. Ceri, "*On the role of statistics in the era of big data: A computer science perspective*" *Statistics & Probability Letters*, 136, 68-72, 2018.

[39] Cfr. audizione dei Proff. Giannotti e Pedreschi (5 dicembre 2017).

[40] Per coda lunga si intendono i prodotti con bassi volumi di vendita che, collettivamente, possono arrivare a rappresentare una quota significativa del fatturato complessivo.

[41] Cfr. audizione Microsoft (9 ottobre 2018). Analoghe considerazioni sono state svolte anche da IBM (22 ottobre 2018).

[42] Cfr. audizione Microsoft (9 ottobre 2018).

a disporre di enormi quantità di dati, si distinguono per cultura e propensione all'investimento[43], e dunque sono stati i primi a sviluppare gli algoritmi capaci di analizzare grandi volumi di dati e tuttora "*innovano e migliorano costantemente la loro capacità di data analytics cercando e acquistando soluzioni computazionali efficienti, risorse umane di eccellenza nonché intere start up innovative*"[44].

Infine, per quanto riguarda le soluzioni per la memorizzazione e l'elaborazione, indispensabili per adottare l'approccio *data driven*, sembra doversi escludere che i soggetti che non ne dispongono al proprio interno versino in una condizione di svantaggio competitivo, considerata la possibilità di acquisire in *outsourcing* i servizi di *cloud computing*, che rende i costi per l'acquisto della capacità di stoccaggio e delle infrastrutture di calcolo sostanzialmente lineari rispetto alle dimensioni dell'attività svolta[45]. Nondimeno occorre considerare che, anche alla luce di quanto emerso nel corso delle audizioni, i soggetti dai quali è allo stato possibile acquisire i suddetti servizi sono proprio i grandi operatori dell'ICT, quali Google, Amazon, Microsoft e IBM.

1.3.3. L'interpretazione e l'utilizzo dei *Big Data*

La disponibilità di informazioni estratte mediante l'analisi dei *Big Data* ha reso possibile un cambio di paradigma anche nel processo decisionale delle imprese, anch'esso guidato dai dati (*data driven decision making*), nel senso che le decisioni possono essere prese direttamente sulla base dei dati, nonché delle correlazioni tra di essi, senza la necessità di una compiuta preliminare comprensione del fenomeno oggetto dell'intervento. In altri termini, in una prospettiva di utilizzo commerciale dei dati, dapprima interviene l'analisi dei fatti, quindi l'azione e infine, e solo eventualmente, la comprensione del fenomeno. Ad esempio, un operatore della grande distribuzione può modificare il posizionamento a scaffale dei prodotti nei propri negozi semplicemente sulla base di correlazioni tra dati, senza bisogno di comprendere le ragioni per le quali il diverso posizionamento ha un impatto positivo sui ricavi di vendita[46]. Laddove invece l'utilizzo dei *Big Data* abbia delle finalità diverse da quelle commerciali, ad esempio nella ricerca medico-scientifica, alle possibilità offerte dalla profilazione non potrà che affiancarsi anche l'apporto del tradizionale metodo scientifico.

Secondo questo nuovo approccio, la disponibilità dei dati assume una valenza di gran lunga superiore rispetto a quella dei modelli interpretativi, giacché, a partire da una grande e variegata mole di dati, gli algoritmi di intelligenza artificiale sono in grado di individuare complessi schemi di relazioni che possono sfuggire ai ricercatori (umani).

Non solo, ma le decisioni in tal modo adottate possono poi essere monitorate ed analizzate con l'ausilio dei dati, dando così vita ad un processo iterativo ed esponenziale in cui i dati sulle esperienze passate forniscono sia *feedback* per il miglioramento di quelle successive, sia *input* per l'adozione di nuove decisioni in ambiti differenti. In tal modo gli algoritmi di *machine learning* possono arrivare a svolgere un gran numero di compiti che in passato richiedevano l'intervento dell'uomo, come ad esempio guidare un'automobile, sulla base dei dati raccolti da tutti i sensori del veicolo stesso, nonché delle correlate applicazioni di analisi dei percorsi stradali e del traffico.

---

[43] Secondo dati PwC, Amazon, Google, Microsoft ed Apple rientrano tra le prime 10 imprese per spesa complessiva in ricerca, mentre Facebook si colloca a ridosso di questo primo gruppo (in 14a posizione). Cfr. https://www.strategyand.pwc.com/innovation1000.
[44] Cfr. audizione del prof. De Streel (19 febbraio 2018).
[45] Cfr. audizione del prof. Gambaro (18 dicembre 2017).
[46] Cfr. OCSE (2015), *Data-driven innovation: big Data for growth and well-being*.

Al di là degli esempi più eclatanti, quali il menzionato progetto di auto a guida autonoma, il descritto approccio decisionale trova applicazione in svariati settori economici, accrescendone la capacità di generare innovazione, sia di prodotto che di processo, innovazione che in ultima analisi risulta anch'essa guidata dai dati (*data driven innovation*).

In generale, nonostante le differenti forme di utilizzo e la rilevanza dei *Big Data* nei diversi settori economici, è possibile ricondurre le principali applicazioni economiche di tale risorsa ai seguenti aspetti.

In primo luogo, i *Big Data* possono contribuire **all'efficientamento e al miglioramento dei processi direzionali, gestionali e operativi** delle organizzazioni. Infatti, grazie alla raccolta e all'elaborazione di dati relativi ai processi interni e al loro monitoraggio è possibile individuare i punti di scarsa produttività e intervenire per migliorare quest'ultima.

In secondo luogo, i *Big Data* possono essere utilizzati per **offrire prodotti e servizi innovativi**, che non potrebbero essere altrimenti realizzati. Si pensi, ad esempio, ai servizi che offrono informazioni agli utenti in merito alle condizioni del traffico sulle arterie stradali, realizzati attraverso la raccolta e l'analisi dei dati di posizione e di spostamento di milioni di singoli utenti.

In terzo luogo, i *Big Data* possono consentire alle imprese di ottenere una **conoscenza altamente dettagliata dei singoli consumatori**, ossia dei loro bisogni e delle loro preferenze. Tale conoscenza può essere utilizzata dalle imprese per realizzare un'elevata **personalizzazione dei prodotti e dei servizi offerti**, aspetto di particolare rilevanza nella fornitura di servizi quali la pubblicità *online* e il commercio elettronico. La comunicazione pubblicitaria *online* si fonda, infatti, sulla capacità delle imprese di offrire agli inserzionisti pubblicitari la possibilità di raggiungere specifici individui, utilizzando nuove modalità negoziali e di allocazione degli spazi che consentono transazioni automatizzate e in tempo reale. Simile è l'utilizzo dei *Big Data* da parte delle piattaforme che distribuiscono contenuti digitali o di *e-commerce*, che possono proporre ai propri utenti beni e servizi in linea con le preferenze individuali. Ad esempio, attraverso l'acquisizione dei *Big Data* personali e relativi alle abitudini del consumatore, alcune piattaforme *online* implementano tecniche di *search discrimination*[47], ossia personalizzano la visualizzazione dei risultati di ricerca *online*.

L'offerta di servizi altamente personalizzati può avere implicazioni molto diverse in funzione dello specifico settore interessato. Nell'offerta di beni e servizi *on line*, la disponibilità di dati che consentono una profilazione dettagliata dei singoli consumatori, può rendere possibile una **differenziazione per singolo utente dei prezzi di beni e servizi**. Per altro verso, nel settore dell'editoria, i *Big Data* rendono possibile un elevato livello di personalizzazione del consumo di contenuti editoriali. Se, da un lato, ciò consente a ciascun utente di avere agevolmente accesso ai contenuti di maggior interesse, dall'altro lato, intensifica fenomeni di cd. *confirmation bias*, per cui gli individui tendono a restare nell'ambito delle convinzioni acquisite, ed *echo chambers*, ovvero di amplificazione dei messaggi, portando ad una polarizzazione delle posizioni, nonché a rischi per pluralismo informativo in ragione del fatto che un dato contenuto e/o prodotto editoriale non viene (tendenzialmente) proposto al di fuori del gruppo di utenti che, secondo il profilo di appartenenza, può *a priori* ritenersi interessato.

[47] L. Chen, R. Ma, A. Hannák, C. Wilson. "*Investigating the Impact of Gender on Rank in Resume Search Engines*", In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (p. 651). *ACM*, 2018. http://personalization.ccs.neu.edu/static/pdf/chen-chi18.pdf

Sotto un diverso profilo, la capacità di acquisire informazioni in tempo reale sul proprio contesto competitivo può consentire alle imprese di modificare e adattare i propri prezzi (*online*) con grande velocità, utilizzando *software* appositi e ricorrendo a regole decisionali predefinite o ad o algoritmi, anche complessi, di *self-learning* che, come già anticipato, potrebbero "imparare" a prendere decisioni in relativa autonomia.

I *Big Data*, infine, trovano un ulteriore campo di applicazione **nell'offerta di nuovi servizi pubblici** contribuendo a migliorare la qualità della vita della collettività. Si pensi, per esempio, al traffico che può essere monitorato costantemente grazie alla condivisione dei dati relativi agli spostamenti registrati dagli *smartphone*, e gestito, nella misura in cui molteplici applicazioni di navigazione satellitare usano questi dati per suggerire ai propri consumatori il percorso più breve e meno congestionato. Ancora, secondo l'approccio *Big Data* possono essere monitorati i tempi di attesa presso gli sportelli pubblici. Analogamente, in ambito sanitario, grazie ai progressi nelle tecnologie di nuova generazione che hanno portato ad una disponibilità crescente di dati biomedici, sono state create banche dati ad accesso libero contenenti dati genomici e clinici di pazienti in forma anonima. Tali database contenenti un gran numero di dati eterogenei costituiscono un grande opportunità per gli scienziati, i quali, avvalendosi di tecniche di analisi dei *Big Data*, possono estrarre nuova conoscenza in maniera automatizzata su una determinata patologia.

Anche le istituzioni pubbliche possono migliorare la propria capacità di azione facendo leva sulla quantità e varietà di informazioni riguardanti le preferenze e le scelte degli agenti economici[48]. Processati attraverso algoritmi di *machine learning*, tali dati, insieme a quelli tradizionali, possono essere impiegati per costruire indicatori dell'attività economica più accurati e tempestivi, ad esempio per stimare il tasso di disoccupazione o il tasso di inflazione, per migliorare le previsioni di variabili rilevanti a fini di *policy*, per misurare il clima di fiducia di consumatori e imprese[49].

La capacità di creare valore attraverso la raccolta e l'analisi di ingenti moli di dati non è un aspetto limitato all'attività delle piattaforme *online*, ma costituisce una potenziale fonte di vantaggio competitivo anche in settori tradizionali, in particolare in quelli caratterizzati da rilevanti asimmetrie informative, nei quali importanti guadagni di efficienza possono derivare dall'elaborazione dei *Big Data*, spesso acquisiti anche attraverso un'attività *offline*.

Prospettiva nell'applicazione dei Big Data nei settori finanziario e assicurativo

Gli approfondimenti svolti nell'ambito dell'Indagine hanno fornito diversi spunti di riflessione in ordine al possibile utilizzo di tecnologie *Big Data* nei settori bancario-creditizio e assicurativo.

Gli operatori del settore bancario-creditizio hanno rappresentato come l'approccio alle tecnologie *Big Data* debba essere necessariamente "prudenziale" in quanto, allo stato, non è ancora sufficientemente chiaro se, a fronte di ingenti investimenti, vi possa essere un effettivo e concreto ritorno economico. Ciò in quanto, da un lato, è estremamente difficile definirne le reali potenzialità e, quindi, le concrete applicazioni operative, considerato che i *Big Data* sono, *prima facie,* dati di qualità mediocre (almeno rispetto ai dati, per così dire "raffinati" di cui dispongono gli istituti di credito o i gestori dei sistemi di informazioni creditizie), la cui effettiva "lettura" e analisi richiedono professionalità specifiche

[48] Sul punto si veda ad esempio l'intervento di apertura al Workshop "*Harnessing Big Data & Machine Learning Technologies for Central Banks*" del Vice Direttore Generale della Banca d'Italia Fabio Panetta, Roma, 26 maggio 2018.
[49] Si vedano ad esempio Daas e Puts (2014), *Social Media Sentiment and Consumer Confidence* e Goolsbee e Klenow (2018), *Internet Rising, Prices Falling: Measuring Inflation in a World of E-Commerce*.

(peraltro ancora difficilmente reperibili sul mercato); dall'altro, e questo rappresenta un ostacolo ben maggiore, nel nostro Paese, i settori in questione operano in un quadro normativo-regolamentare rigoroso e puntuale che non consente alle imprese interessate di effettuare sperimentazioni circa l'efficacia, l'efficienza e l'attendibilità di tali tecnologie innovative. In questo senso si sono espresse le più importanti società[50] sottolineando come in Italia, differentemente da quanto accade in altri contesti europei, non sia possibile utilizzare i cosiddetti "dati alternativi"[51], né i cosiddetti "dati *social*" ai fini della valutazione del merito creditizio; peraltro, gli stessi dati dei soggetti censiti nei sistemi di informazioni creditizie devono essere trattati nel rispetto delle regole contenute nel Codice di condotta[52] recentemente approvato dal Garante per la protezione dei dati personali su proposta delle associazioni di categoria, il cui contenuto tiene conto, almeno in parte, delle nuove sfide poste dalla *digital economy*[53]. Non c'è dubbio, infatti, che negli ultimi anni sono stati sviluppati nuovi prodotti e servizi in ambito *fintech*, quali ad esempio servizi di consulenza *robo-advisor* o servizi di pagamento digitale, la cui diffusione si incrementerà alla luce della c.d. PSD2[54.] Tali nuovi "bacini di informazioni" potranno alimentare nuove elaborazioni basate su tecnologie *Big Data*.

Nel settore assicurativo, gli operatori, nel manifestare innanzitutto un certo disagio per il fatto di ritenersi esposti, a breve, alla pressione competitiva di "giganti del *web*" come Amazon (che sta pianificando, appunto, il suo ingresso nei mercati assicurativi), hanno rappresentato come la diffusione delle tecnologie basate sui *Big Data* lasci intravedere importanti prospettive sia nello sviluppo di forme più competitive di erogazione dei servizi assicurativi sia ai fini dell'ottimizzazione ed efficientamento dei processi interni e della gestione delle polizze.

L'elaborazione dei *Big Data* consentirà infatti alle compagnie assicurative di formulare offerte incentrate sulle caratteristiche sempre più personali della clientela[55], con importanti effetti sulla prevenzione e la riduzione dei rischi, nonché sul miglioramento della gestione e della ricostruzione dei sinistri e nel contrasto alle frodi. Fino ad oggi, tuttavia, l'uso delle nuove tecnologie si è concretizzato principalmente nella diffusione di polizze *online*, anche attraverso il ricorso a siti comparatori, nonché nell'offerta di sconti sulle polizze RC auto vincolati all'installazione delle c.d. "scatole nere". Al riguardo si segnala che, da diversi anni, la materia delle "*black box*" è oggetto di attenzione da parte del legislatore nel tentativo di delineare un quadro normativo-regolamentare che, tuttavia, non ha ancora trovato una compiuta definizione; lo stesso Garante, in occasione dei pareri

[50] Cfr.audizioni di Crif S.p.a.(18 dicembre 2017) ed Experian S.p.a. (28 novembre 2017)

[51] Per "dati alternativi" si intendono i dati relativi al comportamento nei pagamenti ovvero, più in generale, al comportamento economico finanziario di famiglie ed imprese, la cui disponibilità consentirebbe di avere un quadro più esaustivo del soggetto che accede al credito con conseguente abbassamento del livello di rischio per gli istituti di credito.

[52] Trattasi del *"Codice di condotta per i sistemi di informazione gestiti da soggetti privati in tema di crediti al consumo, affidabilità e puntualità nei pagamenti*" approvato dal Garante il 12 settembre 2019 (doc. web 9141941) dopo un complesso lavoro di revisione del vecchio Codice deontologico reso inattuale dalle novità introdotte in materia di protezione dei dati personali dal Regolamento (UE) 2016/679. Le nuove regole, nel disciplinare, tra l'altro, i termini di conservazione dei dati censiti e interrogabili *on line*, prevedono una più consistente profondità temporale (10 anni) per la conservazione dei c.d. dati *off line*; questi potranno essere utilizzati per analisi statistiche e per la costruzione di modelli predettivi, ma soltanto attraverso "*opportune tecniche di cripting o pseudoanonimizzazione*".

[53] Le nuove regole per l'analisi del rischio riguardano infatti non solo i dati su prestiti e mutui, ma anche quelli relativi alle diverse forme di *leasing*, al noleggio a lungo termine e alle più innovative forme di prestito tra privati gestite tramite piattaforme tecnologiche Fintech.

[54] Direttiva (EU) 2015/2366 sui servizi di pagamento nel mercato interno (PSD2).

[55] I clienti (o potenziali clienti) possono essere classificati con maggiore accuratezza sia in base alle loro caratteristiche (che riflettono i diversi livelli di rischio cui gli stessi sono soggetti) sia in base ai loro comportamenti (di guida e non), nonché attraverso la combinazione di modelli "tradizionali" (basati su calcoli di convenienza statistica) con analisi in chiave comportamentale e predittiva (attraverso il ricorso ad algoritmi).

formulati ai sensi dell'art. 154, comma 4, del d.lgs. n. 196/2003 (Codice in materia di protezione dei dati personali), ha rilevato forti criticità sotto il profilo della protezione dei dati personali[56].

Nel corso dell'Indagine è inoltre emerso che, nei prossimi anni, una grande quantità di informazioni sarà detenuta, più che dalle imprese assicurative, dalle imprese produttrici di automobili; si stima infatti che, grazie allo sviluppo della *internet of things*, entro i prossimi 5 anni oltre il 70% delle automobili saranno connesse, con conseguente produzione e raccolta di un'ampissima quantità di informazioni sugli autoveicoli e sui relativi proprietari e/o guidatori da parte delle imprese produttrici che diventeranno dei *gatekeeper* nei loro settori specializzati; è evidente che tali dati potrebbero essere molto importanti per le società che forniscono prodotti assicurativi, tant'è che le compagnie assicurative si stanno impegnando a livello europeo in una consistente attività di *lobbying* finalizzata ad avere accesso al mercato dei dati prodotti dall'industria automobilistica. Al riguardo, sono state evidenziate diverse problematicità, tra cui i profili relativi alla proprietà dei dati raccolti e la possibilità per il singolo automobilista di manifestare concretamente la propria volontà al riguardo.

1.4. Alcuni dati sulla diffusione dell'utilizzo dei *Big Data* nell'economia

Secondo l'*Osservatorio Big Data Analytics & Business Intelligence* del Politecnico di Milano[57], i *Big Data* hanno raggiunto un valore complessivo di 1,4 miliardi di euro nel 2018. Negli ultimi tre anni, il loro valore è cresciuto annualmente in media del 21%. Ad investire sono soprattutto le grandi imprese, che coprono l'88% della spesa complessiva, mentre le piccole e medie imprese rappresentano il 12% del valore. Il rapporto dell'*Osservatorio*, inoltre, ripartisce la suddivisione della spesa in *Big Data* Analytics tra i vari settori merceologici, indicando come primo il settore bancario (28% della spesa), seguito dal comparto manifatturiero (25%) e dal settore telecomunicazioni e media (14%), mentre il restante è coperto da servizi (8%), grande distribuzione (7%), assicurazioni (6%), *utility* (6%) e pubblica amministrazione e sanità (6%).

In termini di impieghi, il rapporto dell'*Osservatorio* indica che il 45% della spesa è dedicata ai *software* (*database* e strumenti per acquisire, elaborare, visualizzare e analizzare i dati, applicativi per specifici processi aziendali), che costituiscono anche l'ambito con la crescita più elevata (+37%). A seguire, i servizi (quali personalizzazione dei *software*, integrazione con i sistemi informativi aziendali, consulenza di riprogettazione dei processi) e le risorse infrastrutturali (capacità di calcolo, *server* e *storage* da impiegare nella creazione di servizi di analisi) coprono rispettivamente il 34% e il 21% della spesa complessiva.

[56] Si veda Relazione annuale 2014, cap. 3 .1, lett. d), nonché, da ultimo, la nota del Presidente dell'Autorità del 1 luglio 2015 in merito all'art. 8 del disegno di legge in materia di concorrenza (AC 3012), laddove il Garante, nell'esprimere apprezzamento per l'avvenuto superamento della precedente previsione normativa (art. 8 d.l. n. 145/2013) sul "servizio unico di raccolta dei dati", ha altresì condiviso, da un lato, la previsione secondo cui l'interoperabilità e la portabilità degli apparati elettronici sia assicurata dagli operatori del settore (denominati "*providers di telematica assicurativa*") i quali devono gestire i dati sull'attività del veicolo in sicurezza e sulla base dello standard tecnologico comune da definire con decreto del Ministro dello sviluppo economico (allo stato non adottato); dall'altra il divieto di utilizzo dei dispositivi per raccogliere dati ulteriori rispetto a quelli necessari al perseguimento della finalità prevista nonché di rilevare la posizione del veicolo in maniera continuativa o sproporzionata.

[57] Si veda il comunicato stampa dell'Osservatorio, 11/12/2018, disponibile su: https://www.osservatori.net/it_it/osservatori/comunicati-stampa/big-data-analytics-italia-mercato-2018 (visitato il 12 febbraio 2019).

2. Principali considerazioni sulla gestione dei *Big Data* espresse dai soggetti partecipanti

Il fenomeno dei *Big Data*, come già detto, si caratterizza per i seguenti aspetti: la disponibilità di una mole di dati che compongono il c.d. *datasphere*, l'eterogeneità delle fonti sorgenti dei dati e la rapidità con cui essi circolano da un punto di origine ad uno di raccolta. Queste caratteristiche, per loro natura, avvalorano il ruolo attualmente svolto dalle reti di comunicazione elettronica, soprattutto quelle di ultima generazione, le cui prestazioni consentono di soddisfare traguardi sempre più sfidanti.

Nei successivi paragrafi si riporta, in sintesi, quanto emerso dai contributi forniti dagli esperti e dalle società che hanno partecipato alle audizioni in relazione al fenomeno dei *Big Data*, con le conseguenti implicazioni economiche, sociali, politiche e normative. In particolare, si riportano le tematiche relative alle attività di profilazione e di anonimizzazione del dato, all'utilizzo di algoritmi, all'acquisizione del consenso per il trattamento dei dati, alla portabilità, interoperabilità e accesso ai dati, all'attività svolta dalle piattaforme digitali.

2.1. Profilazione, anonimizzazione del dato e algoritmi

Le società che hanno partecipato alle audizioni hanno ripetutamente evidenziato che il valore dei *Big Data* non risiede tanto nella diponibilità di una mole di dati, bensì nella loro qualità. I dati, infatti, una volta organizzati ed elaborati, assumono grande rilievo in relazione alle informazioni che sono in grado di fornire e che possono essere utilizzate per scopi commerciali, sociali o politici. I possessori di tali dati possono, ad esempio, estrarre da essi *trend* di consumo e di comportamento dei singoli soggetti, ottenendo una serie di informazioni finalizzate ad orientare e/o adattare, rispetto ai gusti e alle preferenze espresse dai propri utenti/clienti, le scelte commerciali; con le informazioni così ottenute, è possibile perfino determinare le preferenze preventivamente[58]. La funzione predittiva della profilazione, volta ad anticipare i bisogni degli individui, avviene ricorrendo a tecniche di organizzazione e modellizzazione dei dati raccolti, con l'obiettivo di incidere sulle scelte dei singoli individui, adattandole alla realtà che si vive in un determinato periodo di tempo[59]. Quindi, per valorizzare i dati in possesso, di prassi si ricorre alla profilazione, intesa come l'insieme delle attività di raccolta e di elaborazione dei dati inerenti agli utenti fruitori di un servizio, al fine di segmentarli in gruppi a seconda del comportamento rilevato.

Da tale prospettiva deriva che, come evidenziato durante le audizioni, nell'ambito delle attività di profilazione, finalizzate ad orientare o a determinare preventivamente le scelte di *business*, la conoscenza dell'identità personale dell'utente spesso è marginale rispetto alle informazioni possedute del *retailer* (o da altro soggetto portatore di interessi specifici) e opportunamente elaborate tramite l'utilizzo di algoritmi. In alcuni interventi è stato evidenziato che i soggetti che posseggono una mole di dati sono prevalentemente interessati a conoscere gli usi, i costumi e le preferenze degli

[58] Un soggetto, ad esempio, potrebbe entrare in un negozio per comprare un bene, essere identificato grazie a una delle tante tecnologie presenti sul mercato (come RFID nella fidelity card, lo smartphone, il riconoscimento facciale, ecc.) e in base alla profilazione organizzata dall'algoritmo, vedersi praticato un prezzo differente da quello di un altro cliente, che risponde maggiormente alle sue caratteristiche.

[59] Nelle elezioni americane, ad esempio, un candidato presidente, attraverso le analisi di *sentiment* basate su *Big Data* è stato in grado di ottimizzare i propri discorsi, adattandoli al sentir comune dei cittadini dei luoghi nei quali faceva campagna elettorale. Altre ipotesi: i dati delle previsioni metereologiche o epidemiologiche elaborate per un determinato arco temporale, possono essere correlati con i dati relativi alle vendite per verificare se in un negozio ci sono variazioni nelle vendite a seconda delle previsioni fatte; l'utilizzo dei dati sanitari e relativi a una specifica area cittadina potrebbe consentire di verificare se vi sono variazioni dei prezzi sui farmaci a seconda delle patologie che colpiscono l'area in esame.

utenti/clienti, al fine di tracciare e di definire i c.d. "tipi-ideali", ossia una serie di individui-modello in grado di rappresentare le caratteristiche tipiche di migliaia di persone che, rientrando per le loro peculiarità in quel determinato profilo, realizzeranno con alta probabilità le scelte effettuate dal c.d. "tipo-ideale"[60].

Dall'indagine è chiaramente emerso che colui che ricorre a un'attività di profilazione utilizza, tendenzialmente, dati anonimizzati (profilo sul quale si tornerà al par. 4.8), in quanto riesce a ottenere comunque le informazioni di cui necessita al fine di pianificare le proprie strategie di mercato, oltre a non violare la normativa sulla protezione dei dati personali. Le informazioni sull'identità personale di un utente/cliente, pertanto, sembrerebbero avere meno attrattiva, rispetto alla conoscenza sulle caratterizzazioni dei "tipi ideali", sebbene sia stato empiricamente dimostrato come, a determinate condizioni, sia tecnicamente possibile scoprire l'identità di una persona partendo dai dati generici o da metadati (ad esempio gli orari e i luoghi in cui vengono effettuate le telefonate tra due numerazioni).

Da quanto dichiarato dagli operatori di mercato sembrerebbe che, nell'ambito dei processi aziendali, si proceda con l'implementare politiche volte a standardizzare procedure di anonimizzazione dei dati di identità personale. Tuttavia, come evidenziato più volte in audizione, in presenza di dati anonimizzati è possibile, incrociando una serie di *data base* e organizzando e correlando una mole di dati, non solo individuare circostanziati *target* al fine di creare specifiche categorie sulla base dei differenti identikit emersi[61], ma addirittura risalire all'identità del soggetto[62].

Nelle audizioni è stato rappresentato l'importante ruolo svolto dagli analisti informatici, c.d. "*data scientist*", soprattutto in relazione al complesso compito della costruzione degli algoritmi, indispensabili per ottenere dai dati accumulati le informazioni desiderate. Nell'agire quotidiano ogni individuo lascia tracce di dati che raccolti, memorizzati ed organizzati consentono, grazie all'intelligenza degli algoritmi, elaborati dai *data scientist*, di definire correlazioni tra persone, prodotti e servizi. Le "informazioni desiderate", infatti, in considerazione della vastità dei dati disponibili, della varietà della loro struttura (in termini di contenuto e di formati), nonché della velocità di aggiornamento dei medesimi, risultano essere occultate, tra le infinite correlazioni spurie dei dati raccolti, e come tali devono essere "catturate" per portarle alla luce.

Alcuni esperti, nelle audizioni, hanno specificato che l'analisi effettuata sui Big Data può avvenire tramite due tipologie di algoritmi: 1) gli algoritmi *analytics*, che organizzano le informazioni in modo che sia possibile, a certe condizioni, anticipare le scelte, di acquisto o di voto delle persone, determinandone almeno in parte i comportamenti[63]; 2) gli algoritmi di *machine learning,* che abilitano

[60] Cfr.: Interim report AGCOM, pag. 4 sui modelli psicometrici. Negli USA, ad esempio, una donna trenta-quarantenne, sposata, con due figli e che possiede un'auto BMW, avrà una maggiore propensione ad acquistare una vacanza in una capitale europea in un periodo dell'anno differente rispetto ad un *single*, libero professionista.

[61] In audizione è stato citato il caso di dati non personali riferiti ad aggregati di dati di consumo dell'elettricità e del gas, raccolti tramite contatori, che hanno consentito di individuare intere comunità.

[62] Cfr.: L. Sweeney, Simply Demographics Often Identify People Uniquely. Carnegie Mellon University, Data Privacy Working Paoper 3. Pittsburgh 2000. - La ricercatrice, agli inizi dell'anno 2000, lavorando sulla funzione predittiva nelle scelte del singolo, è riuscita risalire all'identità personale di cittadini del Massachusetts incrociando tutti i dati elettorali e ospedalieri (diffusi, pubblici e anonimizzati) che avevano in comune solo il genere, il codice di avviamento postale e l'età. L'esperimento ha dimostrato che nelle attività di profilazione ha grande rilevanza avere a disposizione una grande quantità e varietà di dati; anche la proliferazione di dati depotenziati e liberamente accessibili può consentire di ricavare ulteriori e diverse informazioni.

[63] Essi, ad esempio, trovano applicazione in quei modelli di connessione tra i dati che hanno lo scopo di ottimizzare la mobilitazione porta-a-porta dei comitati elettorali

un apprendimento automatico da parte di un sistema informatico[64], rafforzandone e sviluppandone significativamente le capacità di interazioni complesse. Con gli algoritmi di apprendimento la macchina, man mano che fa esperienza, diventa sempre più autonoma e in grado di fornire informazioni sempre più precise[65].

Nel corso delle audizioni è stato altresì sottolineato che l'uso degli algoritmi finalizzato a pratiche di *price discrimination* potrebbe condurre ad una revisione di alcuni modelli e istituti giuridici inerenti all'autonomia contrattuale delle parti negoziali, che sembrano entrare in crisi nel contesto delle pratiche basate sul *profiling*. Ad esempio, nella decisione di comprare e/o vendere beni e di trattare sui prezzi, il legislatore o il giudice, salvo eccezioni, non analizza la dinamica inerente ai profili economici dell'accordo. Tuttavia, ove il prezzo sia "deciso" da una macchina, in base all'analisi del profilo individuale o di gruppo, dal punto di vista giuridico la situazione cambia radicalmente, essendo il costo del bene determinato non sulla base di "trattative", bensì in ragione di una classificazione attribuita da un algoritmo, che sfrutta le propensioni individuali e che non è nota all'interessato. Inoltre, è stato evidenziato che le imprese, attraverso gli algoritmi, non solo determinano i prezzi, ma possono anche "aggiustarli" in modo dinamico sulla base delle costanti analisi effettuate da programmi informatici. In definitiva, la disponibilità di grandi lotti di dati, elaborati attraverso gli algoritmi, consente una valorizzazione diretta del patrimonio informativo. Come noto, il prezzo di un prodotto/servizio contiene le voci di costo di cui si compone e, in generale, esso, al fine di pianificare le migliori strategie commerciali, viene determinato tramite elaborazioni algoritmiche anche sulla base di una serie di informazioni inerenti ai comportamenti dei clienti (come l'elasticità della domanda al prezzo, le abitudini di consumo, i bisogni consci e inconsci, etc.).

2.2. Gestione del dato e acquisizione del consenso

Nel corso delle audizioni alcuni esperti della materia hanno sottolineato che, durante le fasi di organizzazione, elaborazione ed analisi dei dati, possono presentarsi, inaspettatamente, informazioni che coinvolgono la libertà di decisione dei singoli e della collettività. Il Regolamento (RGPD) sembrerebbe non essere esaustivo rispetto alla riportata questione, essendo legato all'impostazione giuridica ereditata dalla direttiva 95/46, la cui tutela era indirizzata soprattutto alla finalità dell'uso dei dati e al trattamento degli stessi in relazione al singolo interessato. Nel caso dei *Big Data*, i dati sono di sovente trattati per scopi predeterminati solo in termini generali; le finalità non vengono, in realtà, specificatamente individuate *ex ante*, in ragione dell'emersione delle correlazioni fra i dati solo in fase successiva alla raccolta degli stessi. L'impostazione tradizionale (dato ottenuto direttamente dall'interessato, generalmente sulla base del consenso, o comunque individuando preventivamente le responsabilità al trasferimento) si deve confrontare con un fenomeno del tutto nuovo dell'acquisizione massiva di dati personali e dei "parametri d'uso" che vengono acquisiti, per esempio, tramite le app e il loro sistema di permessi[66].

[64] Ad esempio, inserendo delle immagini dei fiori nel sistema operativo e insegnando alla macchina a riconoscerli e distinguerli.
[65] La macchina, ad esempio, può comunicare se in una foto compare un cane o un gatto, indicando anche la probabilità assegnata alla risposta.
[66] Per un'accurata ricognizione sulle questioni connesse alla configurazione delle app (e tra queste anche quelle relative a modalità corrette di manifestazione del consenso da parte degli utenti delle stesse, si rinvia allo studio condotto da ENISA, *Privacy and data protection in mobile applications. A study on the app development ecosystem and the technical implementation of GDPR*, November 2017.

Per tale criticità è stata altresì proposta una soluzione che potrebbe essere rappresentata dal ricorso al *dynamic consent* sviluppato nel diverso contesto del consenso "medico" del paziente in relazione alle c.d. bio-banche[67], purchè soddisfi i requisiti del RGPD. Secondo tale modello, in un primo momento, si presta un consenso ampio sulla base di un'informativa generale circa le possibili finalità del trattamento e, successivamente, una volta individuata specificatamente la finalità di utilizzo dei dati, si riceve una più puntuale informativa con la richiesta di un nuovo e più specifico consenso al trattamento. L'aspetto negativo di una simile proposta è che il soggetto interessato potrebbe subire eccessive sollecitazioni, che psicologicamente potrebbero indurlo a rispondere in modo disattento, senza prestare attenzione al contenuto dell'informativa. Da qui l'opportunità di esplorare, secondo gli esperti, anche altre soluzioni, maggiormente incentrate sulla correlazione fra il rilascio del consenso per il trattamento dei dati ed i rischi che tale trattamento prevede[68].

2.3. Portabilità dei dati, interoperabilità e accesso ai dati

Come noto, il tema della *data portability* è ora espressamente disciplinato all'art. 20 del RGPD, che regola il diritto dell'interessato di ricevere i dati personali che lo riguardano "*in un formato strutturato, di uso comune e leggibile da dispositivo automatico*", forniti a un titolare del trattamento, e il diritto di trasmettere tali dati ad un altro titolare del trattamento senza impedimenti. Durante le audizioni è stato osservato che, pur essendo la portabilità dei dati fondamentale in termini competitivi, essa, tuttavia, può operare efficientemente solo in contesti in cui la strutturazione del dato è bassa, in quanto una manipolazione dei dati più elevata fa venir meno l'incentivo ad effettuare la migrazione di dati[69]. Inoltre, è stato evidenziato che il fenomeno di *lock-in* sociale, unito ad una contrazione del mercato, mina l'effettivo interesse alla portabilità (si pensi, ad esempio, ad un possibile passaggio di un utente da Facebook a Weibo[70]). Pertanto, *lock-in* tecnologici o sociali potrebbero essere di ostacolo al trasferimento del dato da un portale/piattaforma ad un'altra.

È stato altresì sottolineato come la realizzazione e il funzionamento della interoperabilità tra piattaforme dipenda dalla implementazione delle soluzioni tecniche individuate, rispetto a uno o più *standard*, ritenendo, pertanto, che sia ancora troppo presto per capire se la portabilità dei dati tra le piattaforme, nella sua attuale formulazione del RGPD[71], sarà implementata con successo.

Il tema della portabilità dei dati è strettamente collegato sia al fenomeno dei "*Big Data*", sia a quello di "*open data*". Infatti, per coloro che operano in ambito informatico, i *Big Data* si risolvono essenzialmente nella capacità di collegare più *data set*, da cui gli *open data*. Questi ultimi sono tipicamente messi a disposizione dalle amministrazioni pubbliche. Gli operatori hanno, però,

[67] Si veda al riguardo la più articolata posizione seguita nei vari ordinamenti nazionali illustrata nel *Report of the Expert Group on Dealing with Ethical and Regulatory Challenges of International Biobank Research*, *Biobanks for Europe. A Challenge for Governance*, 2011, realizzato per la Commissione europea.

[68] Come di un oggetto altamente infiammabile viene indicato il rischio/pericolo correlato piuttosto che la mera composizione chimica ed il modo d'impiego, così una simile avvertenza potrebbe essere prevista anche per il trattamento dei dati in relazione alle potenziali conseguenze per l'interessato, creando ad esempio una simulazione di quelli che potrebbero essere gli effetti del consenso al trattamento.

[69] Ad esempio, se ci fossero dei dati strutturati in una maniera specifica che ne migliora la visualizzazione/organizzazione, si perderebbe molto tempo ad effettuare una ri-organizzazione degli stessi su un'altra piattaforma.

[70] Sina Weibo è un sito di microblogging cinese. È un ibrido fra Twitter e Facebook, è uno dei siti più frequentati della Cina, si calcola che più del 30% delle persone che hanno accesso a internet in Cina usi Sina Weibo,

[71] Per un approfondimento si veda P. De Hert et al. *The right to data portability in the RGPD: Towards user-centric interoperability of digital services.*

segnalato che è necessario che vi sia un adeguato livello di consapevolezza sugli usi possibili dei *Big Data*, in relazione alla implementazione pratica degli *open data*.

Per quanto riguarda l'accesso ai dati grezzi (dati non strutturati), durante le audizioni è stato riferito che per i *data set* detenuti dalla pubblica amministrazione sono state introdotte norme che prevedono un preciso percorso di messa a disposizione del pubblico di quantità importanti di dati, al fine di consentirne il riuso; Mentre, per i *data set* detenuti da poche imprese, il RGPD prevede il diritto di accesso ai dati personali e la loro portabilità, ai fini del trasferimento dei *data set* da una piattaforma a un'altra e della possibilità di accedere anche ai metadati.

2.4. Utilizzo dei dati di traffico

Nel corso delle audizioni è stato evidenziato che con il RGPD non sono state apportate modifiche al regime di protezione dei dati personali, cui devono attenersi gli operatori di comunicazione elettronica. Permangono, pertanto, solo in capo ai fornitori di servizi di comunicazione elettronica, e non anche agli Over The Top (di seguito "OTT"), gli ulteriori obblighi specifici per il settore definiti dalla "e-Privacy Directive" 2002/58/CE, di cui è in corso la revisione tramite la Proposta di Regolamento e-Privacy. Al contrario degli OTT, ad oggi le società di comunicazione elettronica e i fornitori di accesso ad Internet devono conservare, con rigide procedure di sicurezza, i tabulati di traffico telefonico da rendere disponibili in caso di richieste da parte dell'Autorità giudiziaria. Il rigore che il legislatore riserva ai dati di traffico riguarda sia la loro conservazione, sia il loro ulteriore utilizzo per altre finalità.

In aggiunta ai citati obblighi, gli operatori di telecomunicazioni, in quanto fornitori di servizi di comunicazione elettronica, sono soggetti ad un insieme di obblighi a tutela dell'utenza previsti dalla normativa europea e nazionale di settore. L'offerta di qualsiasi servizio di comunicazione elettronica è soggetta ad obblighi sulla trasparenza, alla pubblicazione di informazioni, alla stipula dei contratti, alle prestazioni ai fini di giustizia, in linea con quanto sancito dal vigente Codice delle comunicazioni elettroniche[72] a recepimento delle Direttive europee.

Nel corso delle audizioni alcuni soggetti hanno rappresentato che, a differenza degli operatori di telecomunicazione, ai fornitori di servizi ed applicazioni *web*, non essendo *provider* autorizzati alla fornitura di servizi di comunicazione elettronica, non si applicano gli obblighi previsti dalla normativa di settore. La fruizione di servizi ed applicazioni offerti da *provider* OTT -, infatti, spesso avviene senza che l'utente possa esprimersi sul trattamento dei dati personali effettuato dal *provider*; in tal caso, infatti, si presuppone che l'utente dia il consenso a tutte le condizioni di utilizzo, comprendenti anche quelle relative al trattamento dei propri dati personali. L'utente, pertanto, spesso è ignaro non solo di dove risiedano i propri dati, ma anche dell'uso che ne fa l'OTT.

Un esempio, presentato in audizione, è fornito dai c.d. "servizi di localizzazione" erogati dagli OTT tramite apposite applicazioni: sebbene il dato generato dall'utente, e che deve essere utilizzato per erogare il servizio, è lo stesso sia per gli OTT che per le società di telecomunicazione, solo queste ultime, per offrire il suddetto servizio innovativo, non possono servirsi del dato di traffico associato all'utente che lo produce: non gli è consentito, per la normativa sul trattamento dei dati personali, associare il "dato tecnico" al "dato cliente". Il dato di traffico che utilizzano gli OTT per offrire il servizio, però, è lo stesso che servirebbe alle società di telecomunicazione, che per erogare servizi

[72] Cfr.: art. 70, 71 e 96.

innovativi ai clienti e per poter utilizzare i dati di traffico generati da questi servizi, dovrebbero ottenere dagli utenti tanti consensi specifici quanti sono i servizi innovativi disponibili. In mancanza di consensi specifici da parte dei clienti, i servizi innovativi non possono essere erogati, riducendo la gamma di offerte potenzialmente disponibili in capo all'operatore di telecomunicazioni. Diverso è il caso di servizi innovativi come "la domotica", offerti dalle società di telecomunicazione, poiché il dato di traffico risulta essere già associato al medesimo cliente/utente che lo produce. In definitiva, le società di telecomunicazione hanno rappresentato che, avendo maggiori difficoltà, rispetto ad altri soggetti che operano sul mercato, ad offrire servizi innovativi, proprio perché risulta essere molto complesso ottenere consensi specifici per ogni singolo servizio innovativo, non possono lavorare con l'interezza dei dati della rete e conseguentemente non riescono ad ottenere campioni statisticamente significativi. Esse, pertanto, hanno segnalato che i loro clienti sono esclusi dal beneficio del legittimo interesse[73], a differenza di quanto avviene per gli utenti degli operatori OTT.

2.5. Piattaforme digitali: pluralismo dell'informazione e potere di mercato

Oggi le piattaforme informatiche svolgono un nuovo ruolo di facilitazione e di intermediazione tra cliente e venditore; questo, tuttavia, pone un problema di neutralità delle informazioni che vengono recepite dai rivenditori e che vengono utilizzate per la formazione dei prezzi. Dal punto di vista del *business* e del mercato si ha, pertanto, un passaggio da un modello lineare a uno stellare, con una ibridazione dei ruoli.

Il mutamento investe anche il "settore dell'informazione", in quanto il canone dei criteri giornalistici e scientifici (funzione di *gate-keeper*, di critica oggettiva) paiono travolti e messi in crisi dalla rivoluzione digitale e informatica. Nel 2018 è stato rilevato che tra le 10 maggiori imprese che operano nel settore dell'informazione e dell'editoria, le prime posizioni sono occupate da soggetti come Google e Facebook, che non sono nati come società editoriali. La presenza di nuove imprese sembra garantire il massimo pluralismo informativo; tuttavia, secondo gli esperti l'aumento delle fonti e delle informazioni non coincide con la loro qualità, che spesso è bassa o inesistente come nel caso delle *fake-news*. Pertanto, per arginare un fenomeno di disinformazione e/o degrado della notizia è necessario individuare criteri idonei a distinguere l'informazione realmente professionale da quella che non lo è, altrimenti il rischio potrebbe essere quello di considerare qualsiasi post, notizia,

[73] Cfr.: *Parere WP 217 del 9 Aprile 2014, n. 6 sulla nozione di legittimo interesse del titolare del trattamento* adottato dal Gruppo dei Garanti UE. In base a tale parere, per essere considerato "legittimo", (quale base giuridica lecita del trattamento) l'"interesse" del titolare deve basarsi su tutte le seguenti e contestuali condizioni (da comprovare): 1. deve essere legale, cioè previsto da una norma di legge nazionale o europea; 2. deve essere concreto (dunque non una astratta enunciazione di principio di un generico interesse a procedere ad un certo trattamento); 3. deve essere – di conseguenza - documentabile e sufficientemente chiaro da poter essere illustrato e giustificato dal titolare in maniera articolata, onde consentire di svolgere una comparazione pratica tra detto interesse e i diritti e le libertà degli interessati (per verificare quale prevalga); 4. deve essere reale, attuale e non speculativo; 5. il correlato trattamento di dati deve essere realmente e concretamente necessario per perseguire l'interesse legittimo del titolare e quest'ultimo deve considerare se non esistano modalità e mezzi meno invasivi che gli consentano di perseguire comunque le finalità del trattamento e realizzare l'interesse; 6. i diritti e le libertà degli interessati devono risultare non prevalenti sull'interesse legittimo del titolare del trattamento, all'esito di uno specifico *balance test,* e cioè una comparazione in ogni specifico caso/trattamento tra il proprio interesse e i diritti e libertà fondamentali degli interessati; il titolare deve condurre (e documentare ai sensi anche dell'*accountability* del RGPD) tale *balance test* tenendo presente: la natura dell'interesse (es: diritto fondamentale, interesse pubblico, etc), il possibile pregiudizio del titolare del trattamento o di terzi ove non fosse possibile procedere al trattamento; la natura dei dati, lo status dell'interessato (es: minore, lavoratore dipendente, etc), le modalità del trattamento (es: su larga scala, profilazione, comunicazione ad elevato numero di destinatari, etc), la natura dei diritti e delle libertà che nello specifico sarebbero coinvolte dal trattamento; 7. vanno considerate le aspettative dell'interessato; 8. vanno valutati comparativamente gli impatti del trattamento sulla sfera degli interessati in rapporto ai benefici derivanti al titolare dallo svolgimento del trattamento.

immagine come fonte di informazione. Peraltro, non è raro il caso in cui una *fake-news* sia prodotta all'interno di un circuito editoriale "professionale", allo scopo di generare attenzione e quindi traffico, o al fine di ottenere consenso politico o per fini economici, ovvero attirare investimenti pubblicitari grazie al traffico generato[74].

Il potere di mercato raggiunto da alcune grandi piattaforme (c.d. GAFA(M) - Google, Apple, Facebook, Amazon e Microsoft) nella fornitura dei servizi digitali è stato segnalato in audizione come un fenomeno in grado di stravolgere radicalmente le dinamiche concorrenziali di numerosi mercati, anche di quelli dove le stesse piattaforme non sono (ancora) attive. La disponibilità di enormi volumi di dati e la loro capacità di acquisirli, elaborarli e sfruttarli amplifica la gamma dei servizi che possono essere offerti a consumatori e imprese. Nel corso delle audizioni è stato altresì evidenziato come la continua espansione e diversificazione delle attività dei grandi operatori digitali renda difficile ipotizzare una possibile erosione, nel breve termine, di alcune posizioni di dominanza. Ciò anche in considerazione dei rilevanti effetti di rete che caratterizzano talune piattaforme multi-versante e della disponibilità di informazioni dettagliate sul comportamento dei consumatori.

La rivoluzione dell'economia digitale comporta peraltro che la competizione non si svolga più all'interno di un singolo mercato, bensì anche in mercati dove gli operatori digitali non sono ancora attivi ma in cui, grazie alla disponibilità dei *Big Data* e alla capacità di elaborarli, potrebbero agevolmente entrare e rapidamente "dominarli". In alcuni casi, la creazione o il rafforzamento di potere di mercato derivano da fenomeni di crescita esterna: la crescente rilevanza assunta dai *Big Data* in alcuni settori suggerisce di porre particolare attenzione anche alle acquisizioni di natura conglomerale.

3. I *Big Data* nell'ecosistema digitale italiano: considerazioni dell'AGCOM

Il lavoro congiunto delle tre Autorità ha consentito, nell'affrontare le tematiche sui *Big Data*, di mettere a fattor comune le rispettive competenze e conoscenze, beneficiando delle differenti prospettive, quali: la regolamentazione propria dell'AGCOM, gli interventi *antitrust* dell'AGCM e il trattamento e la protezione dei dati personali del Garante. Esistono, infatti, tematiche specifiche che, pur impattando su dinamiche complessive, possono essere meglio affrontate per competenza da ciascuna autorità con i rispettivi strumenti normativi, amministrativi e tecnologici. Al tempo stesso è, importante, stabilire un punto di coordinamento per affrontare aspetti generali e promuovere iniziative comuni finalizzate a suggerire orientamenti, raccomandazioni e buone pratiche in una visione complessiva delle politiche pubbliche in materia di *Big Data*.

In generale, ciò che è emerso dall'Indagine è che i fallimenti di mercato nell'ambito delle piattaforme non sono solo quelli "classici" (potere di mercato, esternalità, ecc.) che agiscono dal lato dell'offerta e della struttura di mercato, ma anche quelli, più recentemente oggetto di studio dell'economia comportamentale (*framing, prominence, self confirmation bias, default-bias* ecc.), che riguardano le dinamiche della domanda. Questi ultimi fattori, ove non opportunamente considerati, possono compromettere il sano funzionamento delle dinamiche concorrenziali con particolare riferimento al ruolo esercitato dalle piattaforme sull'effettiva capacità di scelta dei consumatori.

[74] Ricordiamo il recentissimo caso dell'Ambasciata di Svezia che a maggio 2019 è dovuta intervenire formalmente per smentire numerose affermazioni contenute in un servizio del TG2.

Nell'ambito del perimetro di competenze dell'AGCOM, l'utilizzo dei *Big Data* e il conseguente sviluppo delle piattaforme *online* di dimensione globale - che ne sono le principali beneficiarie - hanno un forte impatto su tutti i settori economici tradizionalmente regolati, ma rivestono una particolare rilevanza sul settore dei servizi media audiovisivi e su quello delle comunicazioni elettroniche.

Nel settore dei servizi audiovisivi, la crescita di piattaforme *online* attive nella produzione, distribuzione e condivisione dei contenuti di informazione ed intrattenimento implica che le attività istituzionali dell'Autorità ne vengano influenzate (in senso lato), con riguardo ai profili relativi alla tutela della concorrenza e del pluralismo informativo nel sistema dei servizi di media audiovisivi e dei mezzi di comunicazione di massa, nonché nel mercato della pubblicità (art. 5 del Testo Unico dei Servizi di Media Audiovisivi e Radiofonici - "TUSMAR"-): la garanzia del pluralismo "esterno", ovvero della distribuzione equa delle risorse economiche e tecniche nel sistema delle comunicazioni (art. 43 TUSMAR), la tutela del pluralismo politico e sociale (legge 22 febbraio 2000, n. 28), la garanzia della correttezza e della completezza dell'informazione in quanto servizio di interesse generale (art. 7 TUSMAR), la tutela della dignità umana contro i messaggi di incitamento all'odio e di discriminazione basati su origini etniche, orientamento sessuale, pratica religiosa (art. 32, comma 5 TUSMAR), la protezione dei consumatori nell'ambito pubblicitario (art. 36-bis/40-bis TUSMAR) e la tutela dei minori (art. 34 TUSMAR). Tali elementi verranno analizzati diffusamente nei paragrafi 3.1, 3.2 e 3.5. Se, da un lato, le competenze dell'AGCOM in tale materia sono tradizionalmente ancorate agli specifici mezzi trasmissivi espressamente richiamati dalla normativa di riferimento, dall'altro le tutele riconosciute agli utenti e alle imprese sul mercato si fondano sul riconoscimento e la protezione di diritti di rango costituzionale e sono finalizzate a prevenire taluni rischi e a contrastare taluni effetti che possono prodursi, anche a seguito dell'evoluzione tecnologica, sui mercati interessati.

Per quanto riguarda il settore delle comunicazioni elettroniche, come meglio spiegato nei paragrafi successivi, lo sviluppo dei *Big Data,* e il loro utilizzo da parte delle piattaforme *online* attive su scala globale, ha avuto sia un impatto indiretto su tale settore, soprattutto in termini di redistribuzione dei ricavi lungo la catena del valore, sia un impatto diretto sull'impiego di alcuni servizi oggetto di crescente sostituzione competitiva, da parte degli utenti, rispetto ai servizi tradizionali della telefonia vocale e la messaggistica tramite SMS, e riconducibile all'affermazione di nuovi modelli di *business* basati sulla raccolta e la valorizzazione dei dati che forniscono servizi di comunicazione bi-direzionali tra gruppi di un numero finito di partecipanti (chiamate vocali tramite protocollo IP, messaggi di posta elettronica, servizi di messaggistica o *chat* di gruppo). Da questo punto di vista, le principali sfide che coinvolgono l'Autorità sono quelle relative alla necessità di inquadrare tale fenomeno competitivo, ormai strutturato, nonché di apprezzarne gli impatti, nell'ambito della definizione dei mercati rilevanti e dell'individuazione di posizioni di significativo potere di mercato (art. 16, 17 e 18 del Codice delle Comunicazioni elettroniche - d.lgs. 259/2003 – "Codice Comunicazioni"), nonché della salvaguardia dei principi di interconnessione da punto a punto della rete (art. 42, comma 2, lettera a) del Codice Comunicazioni) e interoperabilità.[75] L'analisi di questi profili sarà l'oggetto dei paragrafi 3.1, 3.2 e 3.5.

[75] L'Autorità può imporre: a) l'obbligo agli operatori che controllano l'accesso agli utenti finali, compreso, in casi giustificati, e qualora non sia già previsto, l'obbligo di interconnessione delle rispettive reti, nella misura necessaria a *garantire l'interconnessione da punto a punto e valutati i servizi intermedi già resi disponibili*; a-bis) in casi giustificati

Oltre agli aspetti competitivi però, occorre sottolineare che il tema dei *Big Data* avrà anche effetti sull'evoluzione delle stesse reti "fisiche" di comunicazione, che stanno evolvendo verso nuove architetture sempre più performanti. Infatti, le reti mobili di nuova generazione presenteranno, grazie alle innovative tecniche di trasmissione, caratteristiche più performanti in termini di capacità di banda e, grazie alla diffusione di terminali (sensori) di ricezione e trasmissione, potranno generare e raccogliere una quantità di dati molto superiore rispetto a quella odierna, con un impatto significativo sui modelli di *business* che si svilupperanno e sulla redditività delle reti stesse. Da questo punto di vista, l'analisi di come la raccolta, la gestione, la valorizzazione e la *governance* dei *Big Data* impattano sulla redditività futura delle reti di nuova generazione, e quindi sugli incentivi degli operatori di mercato a realizzare investimenti e innovazione in infrastrutture nuove e avanzate, fa parte dei compiti istituzionali dell'AGCOM, conformemente a quanto previsto nell'art. 13, comma 6, lettere c), g), e comma 6-bis, lettera d) del Codice Comunicazioni[76]. Tale analisi sarà oggetto del paragrafo 3.3.

Infine, la trattazione proseguirà, nel paragrafo 3.4, con l'analisi dell'impatto che i *Big Data* stanno avendo in settori industriali non di diretta competenza dell'AGCOM, ma comunque interessanti dal punto di vista delle dinamiche competitive per l'estensivo utilizzo di tecniche di analisi dei dati e algoritmi.

Sotto il profilo prospettico, va in ogni caso sottolineata la novità che la rivoluzione dei *Big Data* genera sulla stessa nozione di 'comunicazioni elettroniche' e sulla sua possibile evoluzione in un contesto regolatorio dinamico volto a ricomprendere, all'interno di quella nozione, anche lo scambio dei dati. Al riguardo, si evidenzia quanto previsto dal nuovo Codice europeo delle comunicazioni elettroniche di cui alla Direttiva UE 2018/1972 dell'11 dicembre 2018. In esso, infatti, viene specificato che il trattamento dei dati personali da parte dei servizi di comunicazione elettronica, sia esso in forma di remunerazione o in altra forma, dovrebbe essere conforme al regolamento (UE) 2016/679 del Parlamento europeo e del Consiglio (par. 15). Inoltre, dal momento che i servizi di comunicazione elettronica sono spesso forniti all'utente finale non solo in cambio di denaro, ma in misura sempre maggiore in cambio della comunicazione di dati personali o di altri dati, il concetto

---

e nella misura necessaria, gli obblighi per *le imprese che controllano l'accesso degli utenti finali, onde rendere interoperabili i propri servizi*.

[76] Il Ministero e l'Autorità, nell'ambito delle rispettive competenze, promuovono gli interessi dei cittadini: a) garantendo a tutti i cittadini un accesso al servizio universale, come definito dal Capo IV del Titolo II; b) garantendo un livello elevato di protezione dei consumatori nei loro rapporti con i fornitori, in particolare predisponendo procedure semplici e poco onerose di risoluzione delle controversie da parte di un organismo indipendente dalle parti in causa; *c) contribuendo a garantire un livello elevato di protezione dei dati personali e della vita privata*; d) promuovendo la diffusione di informazioni chiare, in particolare garantendo la trasparenza delle tariffe e delle condizioni di uso dei servizi di comunicazione elettronica accessibili al pubblico; e) prendendo in considerazione le esigenze degli utenti disabili, di quelli anziani e di quelli che hanno esigenze sociali particolari; f) garantendo il mantenimento dell'integrità e della sicurezza delle reti pubbliche di comunicazione; *g) promuovendo la capacità degli utenti finali di accedere ad informazioni e distribuirle o eseguire applicazioni e servizi di loro scelta*.

6-bis. Il Ministero e l'Autorità, nel perseguire le finalità programmatiche di cui ai commi 4, 5 e 6, applicano, nell'ambito delle rispettive competenze, principi regolamentari obiettivi, trasparenti, non discriminatori e proporzionati: a) promuovendo la prevedibilità regolamentare, garantendo un approccio regolatorio coerente nell'arco di opportuni periodi di revisione; b) garantendo che, in circostanze analoghe, non vi siano discriminazioni nel trattamento delle imprese che forniscono reti e servizi di comunicazione elettronica; c) salvaguardando la concorrenza a vantaggio dei consumatori e promuovendo se del caso la concorrenza basata sulle infrastrutture; d) *promuovendo investimenti efficienti e innovazione in infrastrutture nuove e avanzate*, anche garantendo che qualsiasi obbligo di accesso tenga debito conto del rischio sostenuto dalle imprese e consentendo accordi di cooperazione tra investitori e parti richiedenti accesso, al fine di diversificare il rischio di investimento, assicurando nel contempo la salvaguardia della concorrenza nel mercato e del principio di non discriminazione;

stesso di remunerazione dovrebbe essere opportunamente esteso al fine di ricomprendere anche "le situazioni in cui l'utente finale è esposto a messaggi pubblicitari come condizione per l'accesso al servizio o le situazioni in cui il fornitore del servizio monetizza i dati personali raccolti in conformità del regolamento (UE) 2016/679" (par. 16).

Quest'ultima tematica s'intreccia, inevitabilmente, con l'evoluzione della regolazione della cosiddetta *Net Neutrality*, oggi centrata sul ruolo innanzitutto degli Internet Service Provider (ISP) e, dove rilevanti, dei Content Application Provider (CAP). In una prospettiva regolatoria dinamica nella quale lo scambio e la valorizzazione dei dati assumono rilevanza centrale nel modello di *business*, occorre valutare i possibili effetti discriminatori nell'accesso alla rete e ai servizi, derivanti da formule di *zero rating* implicite connesse alla cessione del dato.

Nell'ultimo paragrafo, il 3.5, saranno invece presentati i più recenti sviluppi normativi a livello europeo aventi ad oggetto i temi trattati negli altri paragrafi.

Infine, si tracceranno possibili evoluzioni dell'intervento dell'Autorità, nel panorama delle normative esistenti, quali ad esempio l'analisi di mercato nel settore della pubblicità *online*, la co-regolazione delle piattaforme di *video-sharing*, la vigilanza delle misure di autoregolazione delle piattaforme on-line in tema di contrasto alla *disinformazione,* alla *malinformazione* e alle espressioni d'odio (*hate speech*), che tengano conto del vantaggio competitivo strutturale delle piattaforme *online* globali e della loro potenziale "speciale responsabilità" sui mercati interessati, derivanti dalla capacità di raccolta, selezione e profilazione del dato.

3.1. *Big Data*, mercato pubblicitario, pluralismo e informazione

Le piattaforme digitali, come noto, hanno un crescente impatto sia sul settore della pubblicità *online*, sia sulla produzione e sul consumo d'informazione. Pertanto, lo sviluppo di questi servizi non implica solamente la necessità di dover valutare questi fenomeni dal punto di vista della concorrenza - analizzando la re-distribuzione e la composizione delle risorse pubblicitarie all'interno del mercato - ma anche da altri profili, quali la tutela del pluralismo politico, sociale e culturale, la salvaguardia dell'informazione come servizio di interesse generale, ivi incluso il diritto dei cittadini ad essere informati, la protezione della dignità delle persone, nel quadro del pieno rispetto della libertà d'espressione e della libertà editoriale.

In termini economici, la disponibilità di dati (personali e non) consente ai possessori di tali informazioni di avere una posizione di decisivo vantaggio competitivo nel mercato della pubblicità *online*,[77] ove l'uso del dato è di vitale importanza nell'offerta di prodotti/servizi (c.d. "uso primario del dato"), consentendo, altresì, di sviluppare e migliorare i prodotti/servizi offerti (c.d. "uso secondario del dato"). Il valore molto elevato di queste informazioni per configurare specifici profili di abitudini di consumo spinge le piattaforme *online* a fare in modo di catturare quanta più attenzione possibile, anche attraverso la promozione e la proposizione di contenuti graditi all'utente. L'insieme dei dati così raccolto viene ulteriormente monetizzato con la cessione a soggetti terzi, in alcuni casi violando le norme di *data protection*, in altri casi facendo un uso dei dati personali tale da spingersi oltre il consenso acquisito dall'utente e violando diritti e libertà individuali. Per quanto riguarda la

[77] Si rammenta che a marzo 2019 la Commissione Europea ha sanzionato (nuovamente) Google per pratiche abusive nell'*online advertising*.

raccolta pubblicitaria, questa viene monetizzata attraverso la vendita di appositi spazi per l'*online advertising* presenti sui siti web, raggiungibili dagli utenti tramite terminali fissi e mobili (con sistemi cc.dd. di *programmatic* e *reservation advertising*).[78]

Con riferimento ai compiti istituzionali dell'AGCOM, il fenomeno della concentrazione dei *Big Data* nelle mani di alcune piattaforme *online* ha innanzitutto un impatto riguardante la creazione di posizioni dominanti all'interno dei mercati che compongono il Sistema Integrato delle Comunicazioni.

A tal proposito, il 18 luglio 2019 è stato avviato un procedimento finalizzato all'individuazione e all'analisi del mercato rilevante, all'accertamento di posizioni dominanti o comunque lesive del pluralismo nel settore della pubblicità *online*.[79] Questo procedimento, che si sviluppa a valle degli approfondimenti condotti in questa e in altre indagini conoscitive svolte in questi anni dall'Autorità, costituisce la prima attività istruttoria che coinvolge direttamente anche le piattaforme *online*, e dunque la relazione tra valorizzazione, raccolta e profilazione del dato a fini commerciali. Un possibile esito potrà risolversi, ove necessario, nella definizione di opportuni rimedi, a eventuali posizioni dominanti, che potranno coinvolgere le varie fasi in cui si articola la pubblicità *online* e nelle quali si genera il vantaggio competitivo, incluse le fasi di raccolta e profilazione del dato.

Oltre agli aspetti concorrenziali, il funzionamento del mercato della raccolta dei dati al fine dello sfruttamento pubblicitario può portare a importanti fenomeni patologici che impattano sugli obiettivi della regolazione e sui compiti istituzionali dell'AGCOM. Da questo punto di vista, come già si è analizzato nell'*Interim Report* dell'Autorità in questa Indagine conoscitiva, i principali problemi identificabili sono quelli relativi alla produzione e alla diffusione di fenomeni di disinformazione o della diffusione di contenuti che non rispettino la dignità umana (discorsi d'odio, promozione di contenuti con messaggi di violenza o di discriminazione sessuale/razziale, ecc.).

Con riferimento al tema della informazione e disinformazione, il presupposto di partenza è che le piattaforme *online,* basate su un modello di *business* fondato sulla raccolta pubblicitaria, puntano a catturare quanto più possibile l'attenzione del consumatore. Una volta conquistata, tali piattaforme mirano a far produrre il maggior numero di «azioni» (e.g.: *like*, *scroll*, *search*, ecc.,) all'utente in modo da immagazzinare quanti più dati possibili e della qualità più elevata. Il risultato di questo meccanismo è la profilazione dell'utente funzionale ad una proposizione selettiva di contenuti personalizzati a forte impatto emotivo e fortemente collegati alla propria "storia" di attività *online*. In questo contesto, hanno origine fenomeni noti come *filter bubble* e *self-confirmation bias*, caratterizzati da un meccanismo di causazione circolare per il quale l'utente con le proprie scelte, rivela le informazioni che lo interessano e, a sua volta, la selezione delle informazioni operata dall'algoritmo influenza le scelte dell'utente, confermandone la visione pregressa del mondo. Si realizza in questo modo una ri-proposizione circolare di contenuti confermativi delle proprie opinioni, credenze o convinzioni al singolo utente. Questo fenomeno viene poi ulteriormente rafforzato se l'utente - come accade nella fruizione di contenuti sui *social network* - è immerso in una determinata

[78] *Programmatic advertising*: tecnica di offerta pubblicitaria basata su *software* per l'acquisto e la proposizione di pubblicità digitale, in contrapposizione al processo tradizionale (con negoziazioni tra umani e ordini di inserimento manuali). *Reservation advertising*: acquisto preventivo di *impression* su una data piattaforma da cui verranno distribuiti i contenuti.

[79] Delibera n. 356/19/CONS "Avvio del procedimento volto all'individuazione del mercato rilevante nonché all' accertamento di posizioni dominanti o comunque lesive del pluralismo nel settore della pubblicità on line, ai sensi dell' art. 43, comma 2, del decreto legislativo 31 luglio 2005, n. 177".

realtà di selezione di contenuti anche a seguito dell'interazione con i propri contatti e della conseguente costruzione di *cluster* omogenei di 'amici' che condividono la stessa visione relativa ad argomenti sociali, politici e di attualità. Questo dà origine a fenomeni patologici quali il pensiero di gruppo (*groupthink*) e le "camere d'eco" che agiscono da fattore di polarizzazione, isolando gli ambienti di discussione e predisponendo così ambiti particolarmente favorevoli per coloro i quali vogliano realizzare strategie di *microtargeting* per campagne di *disinformazione* e *malinformazione* aventi ad oggetto tematiche ad alto impatto politico, culturale o commerciale. Tali campagne permettono - attraverso la raccolta e lo sfruttamento di quantità massive di dati, nonché facendo uso di algoritmi, *bot*[80] e account fittizi/anonimi - di influenzare attivamente gli elettori su tematiche di ampio respiro, determinando l'*agenda setting* del confronto politico e aventi poi ricadute su importanti appuntamenti elettorali e quindi sul funzionamento democratico della società.

In ambito nazionale, AGCOM ha istituito un Tavolo Tecnico «*fake news*» con la Delibera n. 423/17/CONS per definire misure di auto-regolamentazione utilizzabili dalle piattaforme, in ottemperanza a quanto disposto dall'art. 7, comma 2, lett. e) del TUSMAR, che recita «*l'assoluto divieto di utilizzare metodologie e tecniche capaci di manipolare in maniera non riconoscibile allo spettatore il contenuto delle informazioni*». A livello europeo, nell'Ottobre del 2018, è stato firmato da alcune grandi piattaforme (Facebook, Google, Twitter, Mozilla), nonché da alcuni importanti attori dell'industria pubblicitaria, un documento, il *Code Of Practice on Disinformation*, contenente alcune misure di auto-regolamentazione che hanno prodotto i primi risultati. Le recentissime analisi preliminari pubblicate dalla Commissione europea sui primi rapporti da parte delle aziende mostrano che in molte aree di intervento, tra cui la riduzione degli incentivi monetari alla disinformazione, la trasparenza dei flussi finanziari della propaganda politica su *Internet* e la lotta agli account falsi, l'auto-regolazione da parte delle piattaforme può risultare, in ultima analisi, insufficiente[81].

a) Le iniziative Agcom sulla disinformazione *online*

Il Tavolo tecnico per la garanzia del pluralismo e della correttezza dell'informazione sulle piattaforme digitali è stato avviato dall'Autorità al fine di promuovere una maggiore comprensione del fenomeno della disinformazione attraverso la partecipazione attiva degli *stakeholder* del sistema dell'informazione e di introdurre misure di contrasto attraverso l'adozione di strumenti volontari e di autoregolamentazione delle piattaforme. Nell'ultimo anno, il Tavolo ha svolto le proprie finalità istituzionali attraverso un'intensa attività di cooperazione e scambio di buone prassi tra i suoi componenti che, ad oggi, ammontano a oltre cinquanta soggetti tra imprese e associazioni nei settori interessati. Le attività svolte sono state condotte in parallelo dai cinque gruppi di lavoro in cui è articolato il Tavolo:

(A) metodologie di classificazione e rilevazione dei fenomeni di disinformazione *online*;

(B) definizione dei sistemi di monitoraggio dei flussi economici pubblicitari, da fonti nazionali ed estere, volti al finanziamento dei contenuti *fake*;

(C) *fact-checking*: organizzazione, tecniche, strumenti ed effetti;

[80] Si tratta *software* che imitano comportamenti umani, come i post su Twitter, o interazioni più ampie.
[81] https://ec.europa.eu/digital-single-market/en/news/annual-self-assessment-reports-signatories-code-practice-disinformation-2019

(D) media e *digital literacy*;

(E) progettazione e realizzazione di campagne informative su disinformazione rivolte ai consumatori.
Segue, pertanto, una sintetica illustrazione dei compiti assolti da ciascun gruppo.

- Il Gruppo di lavoro (A) si è concentrato sull'identificazione di metodi e strumenti per la rilevazione e il monitoraggio dei fenomeni rilevanti di disinformazione *online*, contribuendo alla pubblicazione di diversi rapporti sul sito AGCOM. In particolare, è stato compiuto uno sforzo definitorio fondato su una metodologia sperimentale per la ricognizione e l'analisi quali-quantitativa delle varie distorsioni dell'informazione *online* oggetto di classificazione (distinte in mis-informazione, malainformazione e disinformazione). In tale prospettiva, l'Autorità ha anche istituito l'Osservatorio sulla disinformazione *online*, avviando la sperimentazione di un sistema di monitoraggio della disinformazione *online*, in concomitanza del periodo che precede le elezioni europee di maggio 2019. L'Osservatorio presenta mensilmente i risultati delle elaborazioni svolte a partire da un database di milioni di documenti generati da fonti di informazione e disinformazione, seguendo l'impostazione metodologica già adottata per la redazione del Rapporto "News vs. fake nel sistema dell'informazione".

- Il Gruppo di lavoro (B) ha concentrato la propria attività nella definizione di proposte operative e misure di contrasto alle strategie di disinformazione *online* di tipo commerciale, che realizzano ricavi da pubblicità *online*. Dopo una prima fase, culminata nella predisposizione del rapporto "Le strategie di disinformazione online e la filiera dei contenuti *fake*", sono state avviate due iniziative, aventi ad oggetto l'istituzione di un sistema di monitoraggio della filiera della pubblicità *online* e, in parallelo, la qualificazione dei contesti editoriali di pregio attraverso l'adesione a sistemi di certificazione che rispondono a standard riconosciuti a livello internazionale. Con riferimento al monitoraggio, la finalità perseguita è l'introduzione di strumenti di trasparenza sui soggetti e le attività presenti nella filiera pubblicitaria *online*. Il sistema di monitoraggio progettato a tal fine si compone di un database accessibile a tutti gli operatori del settore, in cui ciascun soggetto indica la fase (o le fasi) della filiera in cui è attivo, specificando i servizi prestati, attraverso quali *brand*, nonché i legami societari/proprietari o di Partnership con altri attori della filiera pubblicitaria. Il sistema è concepito quale naturale evoluzione dell'Informativa Economica di Sistema (IES), che già presenta delle informazioni di natura anagrafica con riferimento agli operatori attivi nei processi di compravendita della pubblicità *online*. Le modalità tecniche di accesso al sistema per i soggetti interessati/attivi nel settore sono in corso di definizione, tenuto conto sia delle norme vigenti in materia di trattamento dei dati, sia dei compiti attribuiti all'Autorità dalle nuove norme in materia di imposta sui servizi digitali (c.d. webtax). Il gruppo di lavoro, infine, ha collaborato congiuntamente al Gruppo C, alla produzione dei contributi presentati dall'Autorità nell'ambito dei lavori della Journalism Trust Initiative (JTI), coordinata dallo *European Committee for Standardization* (CEN) e promossa da Reporters sans frontières (RSF). La definizione ex ante di *Standard qualitativi* appare infatti complementare rispetto all'approccio ex post, che informa la pratica del *fact-checking* come misura di contrasto alla disinformazione.

- Il Gruppo di lavoro C "Fact-checking: organizzazione, tecniche, strumenti ed effetti" ha svolto attività di analisi e studio volte a definire il perimetro di azione e gli effetti della pratica del *fact-checking* sulle modalità di consumo di informazione da parte dei cittadini, ponendosi come obiettivo lo sviluppo, sotto l'egida dell'Autorità, di tecniche, strumenti e soluzioni condivise tra i soggetti interessati. In particolare, sulla scorta dell'esperienza internazionale dell'International

Fact-Checking Network, che ha partecipato alla prima riunione del gruppo, è stata proposta una definizione condivisa di *fact-checking* e delineato il quadro di insieme degli attori che offrono questo tipo di servizi a livello nazionale. Tenuto conto anche dei risultati delle ricerche svolte dal Gruppo di lavoro (B) è stata presentata, sotto l'egida dell'Autorità, una soluzione di mercato, declinabile in una piattaforma di *fact-checking* sul modello dell'iniziativa francese denominata CrossCheck, attualmente però non in corso. Parallelamente all'attività di coordinamento di questa iniziativa, la Segreteria Tecnica del Tavolo ha inoltre formulato due richieste di informazioni alle principali piattaforme *online* (Google e Facebook), specificamente dirette a una maggiore comprensione delle modalità di funzionamento dei servizi di *fact-checking* già resi disponibili, e delle condizioni o fattori di contesto che influiscono sull'efficacia di questo strumento quale misura di contrasto alla diffusione di contenuti *fake* attraverso le suddette piattaforme.

- Il Gruppo di lavoro (D), dedicato alla *media literacy*, ha avviato un primo progetto di lungo termine condividendo indirizzi e riforme adottate in ambito europeo. In particolare, è stata data attuazione a quanto previsto dal "*Action plan against disinformation*" approvato dalla Commissione il 5 dicembre 2018, che raccomanda una tempestiva applicazione delle disposizioni rilevanti introdotte dalla direttiva europea 2018/1808 AVMS. Il primo progetto di *media literacy* avviato dal Tavolo si è rivolto al mondo della scuola e, in particolare, agli studenti delle scuole superiori, in considerazione anche del fatto che alcuni di essi sarebbero stati coinvolti per la prima volta in qualità di elettori alle elezioni europee. Sono stati pertanto prodotti due video sul tema della disinformazione *online* che sono stati veicolati attraverso alcune emittenti televisive nazionali e Facebook. I video sono stati pubblicati anche sul sito web di Agcom a questo indirizzo: https://www.agcom.it/disinformazione. Nella stessa pagina Agcom ha anche dedicato al tema diversi contenuti e in particolare: 1) una descrizione delle regole che i giovani devono seguire per evitare e combattere la disinformazione; 2) un collegamento ai siti dei principali *fact.checkers*; 3) una filmografia dedicata al tema; 4) le attività organizzate anche da altre istituzioni o organizzazioni, quali ad esempio i Corecom. Agcom, inoltre, ha supportato anche la formazione sull'alfabetizzazione digitale avviata da Facebook, Memedia (alfabetizzazione mediatica nel mondo dei meme): il programma di formazione è stato ideato e creato da uno dei massimi esperti in Italia e in Europa sul fenomeno della disinformazione *online*: Walter Quattrociocchi, Coordinatore del *Data Science and Complexity Lab* dell'Università Ca' Foscari di Venezia. L'obiettivo era aiutare i giovani a comprendere meglio le notizie false e il fenomeno della disinformazione. L'iniziativa è stata patrocinata dal "Tavolo tecnico" dell'AGCOM e promosso da "Generazioni Connesse", il Safer Internet centre italiano coordinato dal Miur. L'evento è stato realizzato presso il Binario F di Roma, uno spazio creato a Roma da Facebook e messo a disposizione di imprese, famiglie, accademici, ONG ed editori / media al fine di insegnare e apprendere le competenze digitali. Al progetto in particolare hanno partecipato 40 studenti rappresentanti delle scuole romane, con il compito di diventare ambasciatori di questo importante argomento nelle rispettive scuole. Al termine della formazione, gli studenti hanno infatti ricevuto un kit di informazioni e formazione da utilizzare all'interno delle scuole.

- Il gruppo di lavoro (E) ha proseguito la propria attività pianificando misure dirette a favorire la trasparenza e l'informazione dei consumatori sui rischi legati al fenomeno della disinformazione online, promuovendo a tal fine anche la divulgazione dei risultati dell'attività di ricerca svolta dagli altri gruppi di lavoro. In particolare, è stato realizzato un intervento di formazione e sensibilizzazione dei giornalisti con la finalità di fornire una concreta e immediata risposta ad

alcune criticità segnalate nel Report "News vs. Fake nel sistema dell'informazione" curato dal gruppo di lavoro (A). Pertanto, alcuni componenti del Tavolo, tra cui le associazioni di imprese come Centromarca e dei consumatori (Unione Nazionale Consumatori), hanno segnalato, con propri contributi, l'esigenza di sensibilizzare i giornalisti sul tema del danno reputazionale derivante dalle strategie mirate di disinformazione commerciale e sulla conseguente importanza di un'accurata informazione sui temi e le notizie che hanno un maggiore impatto sulle scelte di consumo, nonché sulla polarizzazione degli utenti di piattaforme *online* intorno a campagne denigratorie su determinati prodotti, servizi o *brand*. Pertanto, di concerto con Centromarca e gli ordini regionali dei giornalisti, sono state organizzate tre giornate di formazione a Roma (4 aprile 2019), Milano (10 maggio 2019) e Napoli (13 giugno 2019).

L'Autorità ha inoltre istituito un Osservatorio sulla disinformazione on line[82] che ha prodotto *report mensile* in occasione delle elezioni europee, basato sulla metodologia già sperimentata in occasione delle elezioni politiche del 2018, allorché si era osservato un picco della disinformazione nella prossimità della settimana del voto, avente come temi principali quelli della criminalità e della immigrazione. Un fenomeno assai ridimensionatosi, invece, nel 2019.

Fonte: Agcom, Osservatorio sulla disinformazione on-line n. 4, 2019

Sempre, nel 2019, in occasione delle elezioni europee, sono stati individuati, di concerto con le piattaforme, gli impegni per garantire non solo la parità di accesso dei soggetti politici alle piattaforme digitali durante l'attuale campagna elettorale per le elezioni dei membri del Parlamento europeo spettanti all'Italia, ma anche un maggiore livello di trasparenza verso gli utenti. Gli impegni hanno costituito il secondo intervento di autoregolamentazione per le campagne elettorali promosso dall'Autorità nell'ambito del "Tavolo Tecnico per la garanzia del pluralismo e della correttezza dell'informazione sulle piattaforme digitali". Dopo la prima applicazione delle linee guida durante le elezioni politiche del 2018, le disposizioni in materia di par condicio per le elezioni europee 2019, adottate con la Delibera n. 94/19/CONS, hanno codificato la buona pratica elaborata con l'introduzione del Titolo VI rubricato "Piattaforme per la condivisione di video e social network". Le

[82] Agcom, https://www.agcom.it/osservatori

disposizioni ivi recate hanno richiamato l'opportunità di promuovere l'adozione condivisa di misure di contrasto ai fenomeni di disinformazione online nell'ambito del Tavolo Pluralismo e Piattaforme, nonché l'impegno delle piattaforme nell'adozione di strumenti volontari a garanzia del pluralismo informativo per la campagna elettorale per il Parlamento UE del 2019. L'identificazione delle misure ha tenuto conto degli strumenti già previsti nel Codice di condotta sulla disinformazione *online* sottoscritto dalle piattaforme il 26 settembre 2018 sotto l'egida della Commissione Europea e delle raccomandazioni della Commissione stessa, contenute nel piano di azione europeo sulla disinformazione. Ad esempio, le piattaforme si sono impegnate, con riferimento ai messaggi pubblicitari di natura elettorale, di informare gli utenti delle piattaforme digitali circa la natura di "messaggio elettorale" e l'identità del soggetto politico committente, inserendo tali informazioni direttamente sul messaggio pubblicitario.

Agcom ha, inoltre, proceduto a rendere uniformi le attività di monitoraggio svolte in ambito nazionale con quelle a livello comunitario; in particolare, ha supportato la definizione di un piano di monitoraggio europeo da parte di una task force specificamente istituita dalla Commissione in ambito ERGA (*European Regulators Group for Audiovisual Media Services*), per monitorare l'implementazione del Codice alla luce dell'Action Plan del dicembre 2018, il cui coordinamento è stato affidato proprio ad AGCOM anche in ragione dell'esperienza già maturata in questa materia.

In particolare, ci sono state due fasi di monitoraggio svolte da 16 Autorità Nazionali di Regolamentazione (ANR) presenti nella task force dell'Erga; una prima fase di monitoraggio si è concentrata specificatamente sugli impegni assunti dalle piattaforme in campagna elettorale, con riferimento in particolar modo alla trasparenza dei messaggi politici elettorali durante le europee; una seconda fase ha riguardato tutti gli impegni assunti nel Codice e in generale la complessiva efficacia di questo strumento (*Scrutiny of ad placements*, *Political advertising and issue-based advertising, Integrity of services, Empowering consumers, Empowering the research community*). Si ricorda che i firmatari del Codice sono Facebook, Google, Twitter, Mozilla e Microsoft.

I risultati della prima fase sono stati pubblicati sul sito Erga; i risultati della seconda fase saranno inclusi in un report che Erga pubblicherà a fine gennaio. Il report sarà utilizzato dalla Commissione per decidere sull'efficacia del Codice e sulle eventuali modifiche da apportare.

Si ricorda, infine, che in preparazione di questi sviluppi regolamentari, già nell'ottobre 2018 l'Autorità aveva invitato, con lettere di sollecito, le principali piattaforme che veicolano inserzioni pubblicitarie e altri contenuti di natura politica ed elettorale in Italia (Facebook, Google e Twitter) ad "assumere un ruolo proattivo" nell'adozione di strumenti di autoregolamentazione. Rammentando gli impegni assunti all'interno del Tavolo, e nell'esercizio delle funzioni di moral suasion proprie del regolatore, il richiamo dell'Autorità era volto peraltro a rendere disponibili gli accessi alle API (*Application Programming Interface*) delle piattaforme per consentire l'attività di monitoraggio concordata dal Tavolo. Per quanto riguarda specificamente Facebook la richiesta era stata formulata anche nell'ambito di incontri svolti in seguito alla vicenda di Cambridge Analytica in relazione alla quale l'Autorità, il 20 marzo 2018, aveva formulato una specifica richiesta di informazioni. Nei mesi successivi, le risposte e i dati forniti da Facebook in merito all'impiego di data analytics per finalità di comunicazione politica da parte di soggetti terzi sono state valutate insufficienti ad acquisire un quadro conoscitivo completo e adeguato ai fini dell'accertamento di eventuali violazioni del pluralismo informativo e dei diritti fondamentali collegati a tale principio, cui è preposta l'Autorità. Pertanto, l'evoluzione del quadro legislativo attuale verso una più precisa definizione dei poteri di

intervento di AGCOM in questa materia appare ormai improcrastinabile ai fini di un'efficace tutela del pluralismo sui mezzi di informazione delle piattaforme *online*.

b) Il contrasto all'*hate speech*

Per quanto riguarda il tema dei discorsi o delle espressioni d'odio (cd. *hate speech*), la diffusione di messaggi violenti è spesso il prodotto del sistema delle camere d'eco sopra descritto. Un tema, quello delle espressioni d'odio, che riguarda tutti i media, inclusi quelli tradizionali, ma che appare sempre più un'emergenza nella discussione sui *social network*, in un preoccupante quadro caratterizzato da un incremento delle segnalazioni delle espressioni d'odio e delle aggressioni ad esse connesse. L'Autorità ha riconosciuto la necessità di avviare dei meccanismi di cooperazione con le piattaforme di condivisione video. In particolare, con la Delibera n. 157/19/CONS, l'AGCOM ha approvato un "*Regolamento recante disposizioni in materia di rispetto della dignità umana e del principio di non discriminazione e di contrasto all'hate speech*". Nella Relazione introduttiva alla delibera, l'Autorità ha sottolineato come "nel corso degli ultimi anni" si sia "registrato un crescente e preoccupante acuirsi, nelle trasmissioni televisive di approfondimento informativo e di *infotainment* delle principali emittenti nazionali, del ricorso ad espressioni di discriminazione nei confronti di categorie o gruppi di persone (target) in ragione del loro particolare status economico-sociale, della loro appartenenza etnica, del loro orientamento sessuale o del loro credo religioso. Tali ripetuti episodi riflettono, indubbiamente, i mutamenti registrati nel dibattito politico, economico e sociale, a seguito del manifestarsi di fenomeni di particolare impatto mediatico-culturale e della relativa trasposizione nella cronaca socio-politica e nell'agenda setting delle diverse forze politiche, in confronti sempre più accesi e polarizzati: attacchi terroristici, fenomeni migratori, episodi di criminalità collegati a vario titolo a specifiche origini etniche o specifici orientamenti sessuali, antisemitismo e così via. Di fronte a fenomeni complessi, che richiederebbero letture multilivello, attenzione ai dati e al contesto, separazione tra il merito del dibattito sulle politiche pubbliche e le valutazioni su specifiche caratteristiche personali, si affermano, con forza, chiavi di lettura semplificate, polarizzanti, divisive e perciò stesso fautrici di discriminazione attraverso espressioni d'odio verso gruppi di persone identificate in base a talune caratteristiche comuni".

Oltre ad introdurre specifiche misure cogenti in capo ai fornitori di servizi media audiovisivi per il contrasto ai discorsi d'odio e alle rappresentazioni errate, violente o discriminatorie di alcuni individui o gruppi di individui, prevede anche che l'Autorità promuova, mediante procedure di co-regolamentazione, l'adozione da parte delle piattaforme di condivisione di video (cd *video sharing platforms*) di specifiche misure volte a contrastare la diffusione in rete, e in particolare sui *social network*, di contenuti in violazione dei principi sanciti a tutela della dignità umana e per la rimozione dei contenuti d'odio. La possibilità data dai *social network* di poter produrre, in forma scritta e senza contatto fisico o visivo con l'interessato e talvolta attraverso *account* in forma anonima, contenuti o messaggi di istigazione all'odio trovano, infatti, nelle camere d'eco il meccanismo più efficace per una loro diffusione sulla rete a grandissima velocità, raggiungendo molti contatti senza produrre ricadute in termini di responsabilità per l'autore.

Tali misure dovranno prevedere anche sistemi efficaci di individuazione e segnalazione degli illeciti e dei loro responsabili, e le piattaforme che le adottano dovranno trasmettere all'Autorità dei *report* di monitoraggio effettuato sulle iniziative prese per l'individuazione dei contenuti d'odio *online*, con l'indicazione anche delle modalità operative e dei sistemi di verifica utilizzati.

3.2. *Big Data*, comunicazioni elettroniche e servizi media

L'impatto dei *Big Data* non si esaurisce nell'ambito della raccolta pubblicitaria, ma coinvolge i modelli di *business* tradizionali in maniera più ampia, grazie soprattutto al processo di convergenza che permette ad alcune piattaforme *online* di entrare nel mercato delle comunicazioni elettroniche e/o nei settori dell'editoria dei media. Ad esempio, alcune piattaforme che forniscono servizi di comunicazione interpersonale si pongono in concorrenza con gli operatori tradizionali del settore delle comunicazioni elettroniche, mentre le piattaforme di condivisione di servizi video, i *social network* e le piattaforme di ricerca, contendono pubblico e risorse economiche agli editori e ai *broadcaster* tradizionali.

Da parte degli operatori tradizionali viene invocato, dal punto di vista regolamentare, l'adozione di un approccio di "*level playing field*", ovvero di graduale avvicinamento dei regimi di responsabilità delle piattaforme/OTT a quello degli operatori concorrenti del mondo *offline*. Nel settore delle comunicazioni elettroniche, questa necessità, sempre più evidente, deriva dallo sviluppo del mercato che, certamente per alcuni servizi tra cui quelli di comunicazione personale e di messaggistica arricchita, richiede l'aggiornamento del quadro regolamentare con impatti, tra gli altri, sulla definizione dei mercati rilevanti delle comunicazioni elettroniche, sulla valutazione delle posizioni di significativo potere di mercato e sulla disciplina a tutela dell'interconnessione, al fine di contrastare la creazione di sistemi chiusi.

Inoltre, la necessità di un approccio coerente tra operatori di comunicazioni elettroniche e OTT viene evocato nelle attività che l'AGCOM svolge a salvaguardia di un'*Internet* aperta, sulla base delle competenze attribuite alle Autorità Nazionali di Regolamentazione dal Regolamento UE 2015/2120. In base a tale Regolamento, viene stabilito che gli utenti finali hanno il diritto di accedere a informazioni e contenuti e di diffonderli, nonché di utilizzare e fornire applicazioni e servizi, ed utilizzare apparecchiature terminali di loro scelta, tramite il servizio di accesso a Internet. Tradizionalmente, le disposizioni attuative di questi principi si sono applicate ai fornitori di accesso ad Internet, ma in contesti concorrenziali caratterizzati da forte concentrazione nella detenzione di *Big Data*, ed a fronte dell'evidenza di forme di discriminazione da parte delle piattaforme algoritmiche, potrebbe rendersi necessario adottare alcune misure a garanzia di trasparenza, equità e di neutralità, al fine di tutelare e salvaguardare la natura "aperta" della rete[83].

Gli altri temi di rilevanza regolamentare relativi al rapporto tra piattaforme *online* e gli operatori di comunicazioni elettroniche sono quelli derivanti dall'uso - da parte dei consumatori - di servizi di messaggistica in sostituzione degli SMS e delle chiamate vocali tradizionali, soprattutto nell'ambito della telefonia mobile. Si evidenzia infatti che i citati servizi VoIP e di messaggistica arricchita prevedono necessariamente l'utilizzo di una numerazione telefonica mobile. Questo tipo di applicazioni prevede la preliminare registrazione dell'utente tramite numero telefonico mobile, che lo identifica e lo abilita a inviare messaggi e a effettuare chiamate vocali verso gli altri utenti registrati al medesimo servizio e presenti nella rubrica telefonica personale contenuta nel dispositivo mobile.

Dal punto di vista tecnico, per offrire i servizi sopra descritti, i nuovi *player* devono ricorrere necessariamente, come identificativo, all'uso delle numerazioni telefoniche mobili degli utenti

[83] Su questo tema si veda l'adozione del Regolamento (UE) 2019/1150 che promuove l'equità e la trasparenza per gli utenti commerciali dei servizi di intermediazione online. L'argomento viene trattato più diffusamente nel capitolo 3.5.

registrati, in assenza delle quali non potrebbero consentire l'invio dei messaggi od offrire il servizio voce. Questo tipo di applicazioni, infatti, a differenza dei *social network*, deve poter individuare un singolo utente attraverso la numerazione telefonica, che viene utilizzata al posto del *nickname*. Dal punto di vista regolamentare, questi servizi, che permettono all'utente di comunicare vocalmente o tramite messaggio con altri utenti su qualsiasi supporto mobile, sono del tutto assimilabili al "*Servizio di numero unico o personale*" identificato, dall'articolo 1, lettera m), del Piano di Numerazione Nazionale (PNN), come "*servizio che permette al sottoscrittore di essere raggiunto, tramite uno stesso numero non geografico, ad un insieme discreto di possibili destinazioni*" [84].

Tuttavia, dal punto di vista formale, l'uso *indiretto* delle numerazioni non qualifica tali soggetti come destinatari degli obblighi del PNN, e per questo motivo in Italia l'uso indiretto delle numerazioni è stato disciplinato dalla legge n. 124/2017. Tale legge, all'art. 1, comma 44, istituisce, presso il Ministero dello sviluppo economico, il registro dei soggetti che usano indirettamente risorse nazionali di numerazione, stabilendo che alla sua "tenuta" si provveda "*ai sensi dell'articolo 1, comma 6, lettera a), numero 5), della legge 31 luglio 1997, n. 249*"; disposizione quest'ultima che attribuisce all'AGCOM la tenuta del Registro degli Operatori di Comunicazione (ROC). Tuttavia, se dal punto di vista della trasparenza e degli oneri di comunicazione dell'attività, l'attuale normativa nazionale equipara - di fatto - i fornitori di servizi tramite piattaforme *online*, la mancata qualificazione di questi operatori come fornitori di reti e servizi di comunicazione elettronica – e quindi l'esenzione dall'obbligo di autorizzazione ai sensi dell'art. 25 del Codice Comunicazioni - impedisce l'applicazione, da parte dell'Autorità, di corrispondenti obblighi di accesso e interconnessione (art. 40, 41 e 42 del Codice Comunicazioni) in caso in cui si verificassero rifiuti a contrarre ingiustificati o fallissero delle negoziazioni di interconnessione con le piattaforme *online*. Da questo punto di vista, sono state introdotte alcune novità dal nuovo Codice europeo sulle Comunicazioni elettroniche, come meglio spiegato nel paragrafo 3.5.

Oltre ai temi relativi all'impatto dei nuovi modelli di *business* sulle filiere tradizionali, alla necessità di rafforzare i presidi a garanzia dell'interconnessione e del carattere "aperto" della rete Internet, nel corso dell'Indagine è emersa un'ulteriore criticità relativa alla diversa cornice regolamentare in cui le piattaforme e gli operatori telco operano, ovvero la possibilità di raccogliere e trattare i dati. In particolare, è stato segnalato dagli operatori di comunicazioni elettroniche come essi abbiano limiti più stringenti per il trattamento dei dati rispetto agli OTT – ad esempio riguardo alla storicità dei dati della clientela - laddove questi ultimi hanno sempre la disponibilità del dato immagazzinato, con la possibilità di "richiamarlo" per utilizzarlo in qualunque momento, con evidenti implicazioni positive sulla dinamicità e qualità dei servizi offerti.

Allo stesso modo, anche i rappresentanti dell'editoria e dell'audiovisivo hanno evidenziano la difficoltà nel poter fruire dell'insieme di dati generati dai consumatori quando accedono ai contenuti, anche originali, tramite piattaforme digitali (come Youtube o *social network*), a differenza degli OTT che, invece, nel porre in essere operazioni di condivisione dei contenuti altrui sulle proprie piattaforme, producono dati e metadati[85] che rimangono nella loro esclusiva disponibilità. In assenza di un *level playing field* e di norme consolidate sull'interoperabilità e sulla portabilità dei dati, gli operatori di comunicazione elettronica, gli editori e i fornitori di contenuti tradizionali lamentano

[84] Si veda l'allegato A alla delibera n. 8/15/CIR.

[85] Persino il *login* ai servizi dei fornitori di contenuti tramite le piattaforme social (c.d. *social login*) produce dati che vengono restituiti solo in parte a tali fornitori, a vantaggio della piattaforma.

dunque uno svantaggio competitivo rispetto agli OTT, derivante sia dalla impossibilità di accedere, analizzare e utilizzare, sempre nel rispetto del principio di tutela dell'utente, dati raccolti e generati dai propri utenti, sia dalla difficoltà di acquisire clienti in assenza di procedure strutturate di portabilità dei dati.

3.3. *Big Data* e sviluppo di reti e servizi innovativi (5G, IoT, M2M, AI)

Il tema dei *Big Data* nell'ambito delle comunicazioni elettroniche ha risvolti importanti sui processi di connessione degli oggetti e delle macchine alla rete internet (*Internet of Things* e comunicazione *Machine-to-Machine*), in particolare nella transizione verso le reti mobili di quinta generazione (5G). La trasformazione "digitale" sta infatti interessando tutti i settori dell'industria, con un forte impatto – come già delineato - sugli operatori che forniscono reti e servizi di comunicazioni elettroniche che, da un lato, devono far fronte alle sfide derivanti dall'ingresso di nuovi concorrenti (soprattutto "nativamente digitali") e, dall'altro, non possono farsi sfuggire le opportunità derivanti dall'adozione di modelli operativi altamente flessibili e *cost effective*.

Poiché il livello di investimento degli operatori di comunicazioni elettroniche continua ad essere elevato per lo sviluppo di reti fisse (NGA) e mobili (5G) innovative, diventa ineludibile l'adozione di processi innovativi di utilizzo delle risorse disponibili, a fronte di fattori come l'elevata competizione su servizi e tecnologie, la continua crescita del traffico dati a fronte di ARPU tipicamente in contrazione, la necessità e l'opportunità di sondare nuovi mercati e servizi, nonché quella di soddisfare le crescenti necessità dei clienti in tutte le fasi del ciclo di vita dei prodotti/ servizi.

Le imprese di comunicazione elettronica audite hanno sottolineato come una pianificazione evoluta e a prova di futuro delle reti di comunicazioni elettroniche non possa prescindere dall'utilizzo dei *Big Data & Analytic*s e dalla virtualizzazione della rete. Infatti, mentre i *Big Data* consentono ad un operatore di comunicazioni elettroniche la flessibilità necessaria per dirigere le risorse nei tempi e nei modi più efficaci ed efficienti (ossia, quando e dove necessario) senza i vincoli derivanti dal dispiegamento di piattaforme verticali per servizio, gli *analytics* possono consentire processi predittivi per interventi preventivi e di pianificazione sia a livello di rete che di servizi.

Le audizioni hanno consentito di evidenziare come, all'interno delle società di comunicazioni elettroniche, il ruolo delle infrastrutture e dei servizi di rete a supporto della generazione dei *Big Data* sia duplice. Il primo consiste nell'abilitare la raccolta, il trasporto e la memorizzazione (es. *cloud computing* e *data center*) dei dati prodotti da fonti terze (es. IoT/M2M, *smart metering*, *devices*, *smart cities*, applicazioni e servizi). Il secondo riguarda la generazione autonoma di dati a livello di rete, di applicazione/servizio e di procedure/processi operativi; in questo caso l'effettiva utilizzabilità dei dati e le relative modalità devono essere coerenti con le leggi comunitarie in tema di *privacy* e di protezione dei dati (v. in merito il cap. 4).

In generale, gli operatori di comunicazioni elettroniche inizialmente hanno sviluppato le proprie conoscenze ed esperienze partendo da casi di utilizzo interni dei *Big Data*, e solo successivamente si sono indirizzati verso la monetizzazione esterna, mediante servizi e applicazioni per i mercati verticali. Per quanto riguarda le iniziative di sfruttamento interne da parte degli operatori di comunicazioni elettroniche, le principali aree di intervento sono: a) il *marketing* e l'assistenza ai clienti, laddove la maggior parte delle iniziative è incentrata sull'*intelligence* proattiva dei clienti (riduzione dei bisogni e abbandono dei clienti), sulle campagne/promozioni personalizzate e sul miglioramento della qualità del servizio; b) le *operational performance*, in cui ci si concentra

sull'analisi predittiva per migliorare le prestazioni della rete e la qualità del servizio, automatizzando i processi e ottimizzando la gestione delle risorse. I casi di utilizzo dello sfruttamento esterno dei dati mobili, anonimi e aggregati, sono invece principalmente rivolti ai servizi di trasporto intelligente, *footfall analytics,*[86] *marketing*, pubblicità e *Big Data as a service.*[87]

In questo contesto, oltre al tema relativo all'utilizzo dei *Big Data* è stato segnalato in audizione come la tecnologia 5G estenderà le applicazioni di Intelligenza Artificiale ("IA") in maniera sinergica. Infatti, le caratteristiche di grande capacità di traffico *ultrabroadband*, in grado di espandere ulteriormente l'*Internet of Things*, e gli innumerevoli ambiti di applicazione permetteranno di incrementare sempre di più la raccolta dei dati e gli algoritmi conseguenti. Con l'introduzione del 5G si assisterà, per la prima volta, ad una complessa interazione tra reti molto più performanti e flessibili: aumenterà il numero di celle, si avranno soluzioni di virtualizzazione e dinamicità della rete per l'elaborazione di processi complessi, non gestibili efficientemente con operazioni manuali, che richiederanno nuovi sistemi basati su IA. La gestione di una mole crescente di dati e la flessibilità delle reti 5G consentirà di cogliere le opportunità derivanti dall'uso di tecniche avanzate di IA per un utilizzo efficiente della rete e per una proposizione di servizio adattata in modo dinamico al singolo caso d'uso.

Dal punto di vista operativo, l'utilizzo dei *Big Data*, anche in combinazione con tecniche di IA, può offrire soluzioni di valore nei seguenti ambiti: a) sviluppo della rete: l'analisi congiunta dei dati di qualità, di traffico ed utilizzo, dei disservizi e dei reclami può consentire un'ottimizzazione dei processi di pianificazione e creazione della rete; b) manutenzione proattiva della rete: la raccolta e l'analisi di eventi ed allarmi provenienti dalla rete consentono l'identificazione preventiva di guasti o malfunzionamenti e l'effettuazione degli interventi di manutenzione prima che si verifichino dei disservizi; c) gestione della sicurezza della rete: la raccolta, l'analisi di eventi ed allarmi provenienti dalla rete ed il loro confronto con valori di soglia permettono di individuare e gestire potenziali attacchi alla sicurezza della rete.

La capacità adattiva delle nuove reti, accoppiata alla pervasività dell'*Internet* delle cose, porterà in tempi brevi ad avere centinaia di milioni di apparati (sensori, attuatori e terminali) connessi, con un numero notevole di strumenti in grado di raccogliere e gestire dati. Con il 5G si concretizzerà quindi la possibilità di avere una rete estremamente efficiente che permetterà lo sviluppo di servizi innovativi, tra cui – solo per citarne alcuni - il monitoraggio di aree sensibili o inaccessibili (ad es. a seguito di calamità naturali), la raccolta e comunicazione in tempo reale di dati sulla mobilità, sullo stato di affollamento per grandi eventi, sulla situazione dei rifiuti, sullo stato dell'illuminazione metropolitana, innovazioni quali la c.d. ambulanza connessa, che permetterà la condivisione in tempo reale dei parametri vitali del paziente e la videochiamata ad elevata risoluzione tra il personale dell'ambulanza e il personale medico dell'ospedale in cui verrà portato il paziente.

Con riferimento all'IA, l'8 aprile 2019, il Gruppo di esperti di alto livello, coordinato dalla Commissione Europea, ha pubblicato le proprie Linee Guida per un'etica dell'IA[88]. Secondo tali Linee Guida, un sistema di IA affidabile dovrebbe essere: 1) legale - nel senso di rispettare ed essere

[86] Si intende la raccolta e l'analisi dei dati relativi alla permanenza degli utenti in prossimità dei punti di interesse (es.: negozi, ristoranti, musei).
[87] Ovvero l'offerta alle aziende delle capacità di gestione e analisi di *Big Data*.
[88] https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai

compliant con tutte le leggi e i regolamenti applicabili al settore rilevante; 2) etico, ovvero rispettoso dei princìpi e dei valori morali umani e 3) robusto da un punto di vista tecnico.

In linea generale, le Linee Guida propongono vari requisiti chiave che i sistemi di IA dovrebbero soddisfare per essere considerati affidabili, tra i quali la garanzia della presenza di un controllo e una supervisione umana, il rispetto dei principi di non discriminazione ed equità nella propria relazione con gli esseri umani coinvolti nelle attività, la massimizzazione del benessere sociale e di quello ambientale e, infine, il rispetto del principio di responsabilità, per il quale dovrebbero essere istituiti meccanismi di monitoraggio dei risultati dei sistemi di IA e di individuazione di precisi responsabili per il loro funzionamento.

All'interno di queste Linee Guida sono presenti anche prescrizioni potenzialmente applicabili alla tematica dei *Big Data*. In particolare, si richiama il fatto che le modalità di raccolta dei dati devono garantire il rispetto della *privacy* e della protezione dei dati personali, devono avvenire nell'ambito di adeguati meccanismi di *governance* di tali dati, tenendo conto della qualità e dell'integrità dei *database* e garantendo un accesso legittimo da chi ne faccia ragionevole richiesta. Si raccomanda infine che le modalità operative dei modelli di *business* di IA siano trasparenti, tracciabili e conoscibili per gli esseri umani che vi interagiscono, i quali debbono avere contezza delle capacità e dei limiti del sistema stesso.

Dal punto di vista dell'AGCOM, il monitoraggio della diffusione delle varie tecnologie e applicazioni dell'*Internet of Things*, del *Machine-to-Machine* e dell'Intelligenza artificiale, lo sviluppo dei servizi e delle reti 5G, l'analisi e l'individuazione delle *best practice* sulla *governance* della raccolta e della gestione dei *Big Data* tra i vari attori dell'ecosistema che forniscono questi servizi, saranno campi di attività fondamentali per garantire il successo dello sviluppo delle reti e dei servizi innovativi che insistono sulle infrastrutture di comunicazioni elettroniche.

3.4. *Big Data* e altri settori

L'analisi dell'impatto dei *Big Data* nelle industrie tradizionali non può fermarsi agli effetti sul settore delle comunicazioni elettroniche e delle sue evoluzioni, dell'editoria e dell'audiovisivo, ma deve estendersi ai settori verticali che in maniera più intensiva saranno caratterizzati dall'uso dei dati.

Da questo punto di vista, l'utilizzo di *Big Data* nei mercati contigui a quello delle comunicazioni elettroniche presenta diversi elementi patologici che potrebbero giustificare la necessità di rafforzare i presidi regolatori. Tali elementi patologici possono essere legati sia a profili di carattere generale, relativi, ad esempio, alle modalità di raccolta ed utilizzo dei dati, sia a profili squisitamente legati a dinamiche specifiche di alcuni settori industriali.

Nelle audizioni congiunte sono stati esaminati, tra gli altri, il settore bancario-assicurativo e il settore del *credit reporting* e *data brokering*, di interesse dell'Autorità poiché sono, ad oggi, tra i più interessati ad un uso massiccio dei *Big Data* e degli algoritmi.

Con riferimento ai profili generali, nell'ambito alle problematiche relative alla raccolta dei dati personali alcuni esperti hanno rappresentato la necessità di un rafforzamento delle normative in materia di tutela del consumatore e *privacy* con riguardo al consenso al trattamento dei dati personali. Studi empirici hanno, infatti, mostrato che gli utenti non prestano una diligente attenzione a lunghe informative sulla *privacy* e questo ha come risultato che le transazioni di questo tipo siano soggette a forti asimmetrie informative.

Dati questi problemi, un possibile approccio normativo potrebbe essere quello di valutare l'articolazione delle *privacy policy* sulla base delle caratteristiche e delle esigenze degli utenti, anche valutandone la proporzionalità, cosicché questi possano decidere con maggior consapevolezza in materia di consenso al trattamento dei propri dati.

Ad esempio, nell'ambito dell'*interim report* dell'Indagine pubblicato dall'AGCOM, era stato sottolineato come in molte occasioni l'utente non è in grado di comprendere esattamente se i "permessi" di accesso ai suoi dati, richiestigli dall'applicazione in corso di installazione sul proprio *smartphone*, siano coerenti e proporzionati agli usi cui è destinata l'applicazione stessa. Pertanto, uno degli ambiti di approfondimento ed eventuale intervento potrebbe essere quello di valutare la proporzionalità di una *privacy policy* che preveda l'accesso, da parte - ad esempio - di un'applicazione di fotografia professionale, alla funzionalità del microfono dello *smartphone* sul quale è installata. In aggiunta, nell'ambito dell'*interim report* è stato osservato che il confine tra dati personali e non personali è divenuto più incerto, e che pertanto una regolazione orientata a proteggere i dati da un punto di vista generale piuttosto che per determinate categorie di dati potrebbe risultare più efficace.

Con riferimento alle problematiche specifiche dei singoli ambiti settoriali, il tema maggiormente sottolineato dagli esperti sentiti in audizione è quello relativo all'attuale regime di responsabilità delle attuali piattaforme *online*. In particolare, è stato più volte sottolineato come queste piattaforme, essendo assoggettate solamente alle regole di responsabilità definite nella direttiva sull'*e-commerce* in qualità di *hosting* o *caching providers*, siano prive di un presidio regolatorio efficace. Secondo gli esperti auditi le piattaforme *online* che selezionano, organizzano e promuovono contenuti personalizzati per i propri utenti, utilizzando *Big Data* e strumenti algoritmici dovrebbero essere sottoposti ad un chiaro regime di responsabilità editoriale.

Nel settore delle banche, e in particolare delle assicurazioni, il tema più attuale è quello dell'utilizzo dei *Big Data* per personalizzare/discriminare i singoli profili degli utenti. Le audizioni hanno messo in rilievo l'importanza degli investimenti in tecnologie per la raccolta e l'elaborazione di grandi quantità di dati necessari a questo fine, così come la necessità di formare o avere accesso a risorse umane qualificate, senza le quali le imprese non saranno in grado di trarre vantaggi dall'evoluzione tecnologica in atto e, in caso di entrata in questi settori dei grandi OTT, non saranno in grado di fronteggiarne la concorrenza.

3.5. *Big Data* ed evoluzione del quadro regolamentare europeo

Le problematiche sopra descritte e il fenomeno della centralizzazione dei *Big Data* nelle mani di pochi *player* internazionali di grandi dimensioni sono da tempo sotto l'attenzione delle diverse Autorità di regolamentazione - europee e non - e delle Istituzioni comunitarie, tutte consapevoli della necessità di dover adottare politiche volte a creare un maggior equilibrio trai diversi attori in campo. A tale proposito, sia nella letteratura che nelle proposte di *policy* portate all'attenzione del pubblico, sono state valutate soluzioni di tipo strutturale o comportamentale volte – attraverso strumenti quali l'identificazione di una posizione di dominanza, la previsione di limiti *ex ante* all'integrazione verticale e orizzontale delle piattaforme - a limitare il potere contrattuale delle piattaforme e, contemporaneamente, a fornire alle Autorità e agli utenti opportuni strumenti di *empowerment* nell'ambito del processo di acquisizione, gestione e valorizzazione dei *Big Data* [89].

[89] Ad esempio, negli Stati Uniti la Senatrice E. Warren del Partito Democratico ha annunciato, nel marzo del 2019, la proposta di una legislazione che dispone che le grandi piattaforme tecnologiche – ovvero le aziende con un fatturato

A livello comunitario, si stanno moltiplicando gli sforzi per rafforzare i presidi sulle piattaforme. Ad esempio, il recente documento "*A Union that strives for more: my Agenda for Europe*" della Presidente della Commissione Ursula von der Leyen ha esplicitato la volontà politica del nuovo esecutivo europeo di adottare prossimamente un nuovo *Digital Services Act* che "*aggiornerà le regole di responsabilità e di sicurezza delle piattaforme digitali, dei servizi e dei prodotti digitali al fine di completare il Mercato Unico Digitale*".

Recentemente sono stati già adottati importanti atti finalizzati a rafforzare i presidi regolamentari sulle piattaforme *online*. Vale in questa sede citarne tre: la nuova Direttiva sui Servizi media audiovisivi (Direttiva 2018/1808) approvata il 14 Novembre 2018, che modifica la precedente Direttiva 2010/13/UE; il nuovo Codice europeo sulle Comunicazioni elettroniche (Direttiva 2018/1972 approvata nel mese di Dicembre 2018) e il Regolamento (UE) 2019/1150, approvato il 20 giugno 2019.

Nella Nuova Direttiva sui servizi media viene innanzitutto sottolineata la crescente affermazione di nuovi operatori, fra cui le *piattaforme per la condivisione di video*, che operano nel mondo *online* e si pongono in concorrenza sull'*audience* e la valorizzazione delle risorse pubblicitarie nei confronti dei fornitori di servizi media tradizionali. Alla luce di tale riconoscimento, alcune misure di regolamentazione tipiche del mondo audiovisivo vengono estese anche alle piattaforme *online*. In particolare, nel Considerando 5) della nuova Direttiva, si stabilisce che i servizi di *social media* o *social network* dovrebbero essere sottoposti a regolamentazione se la fornitura di programmi e di video generati dagli utenti costituisce una loro *funzionalità essenziale*. Inoltre, viene stabilito che, benché tali piattaforme *online* non detengano esattamente una responsabilità editoriale sui contenuti veicolati, tuttavia, in genere, determinano l'*organizzazione* di tali contenuti, ossia programmi, video generati dagli utenti e comunicazioni commerciali audiovisive, *anche in modo automatizzato o con algoritmi*.

Pertanto, in base al riconoscimento che tali piattaforme promuovono contenuti di informazione e intrattenimento come "funzionalità essenziale" e che sono predisposte per "organizzare" la visione di contenuti, esse dovrebbero essere tenute ad adottare le misure appropriate per tutelare le varie categorie di consumatori. Secondo l'art. 28-*ter* della nuova Direttiva, tali piattaforme dovrebbero adottare misure per impedire che video generati dagli utenti e comunicazioni commerciali audiovisive possano nuocere allo sviluppo fisico, mentale o morale dei minori, possano istigare alla violenza o all'odio nei confronti di un gruppo di persone o singoli individui, e, infine, dovrebbero proteggere il grande pubblico da programmi, video generati dagli utenti e comunicazioni commerciali audiovisive che includano contenuti la cui diffusione costituisce un'attività che rappresenta un reato.

Le misure sopra descritte dovrebbero essere attuate tramite una strumentazione che può essere sia di tipo tecnico che contrattuale, e attraverso procedure di co-regolamentazione e auto-regolamentazione.

---

annuo globale di 25 miliardi di dollari o più e che offrono al pubblico servizi di intermediazione *online*, uno come un *marketplace* o una piattaforma per il collegamento di terze parti - possano essere designate come *platform utilities* e che debbano essere separate – dal punto di vista strutturale e proprietario - da qualsiasi partecipante/utente commerciale su quella piattaforma. Le *platform utility* dovrebbero quindi soddisfare uno standard di trattamento equo, ragionevole e non discriminatorio nei confronti dei loro clienti e non dovrebbero essere autorizzate a trasferire o condividere dati con terzi. Per le imprese più piccole – ovvero quelle con un fatturato annuo globale compreso tra 90 milioni di dollari e 25 miliardi di dollari – non sarebbe prevista la separazione strutturale e proprietaria, ma le loro piattaforme sarebbero tenute a soddisfare lo stesso standard di trattamento equo, ragionevole e non discriminatorio richiesto a quelle più grandi.

In particolare, la Direttiva stabilisce che l'adeguatezza delle misure tecnico/contrattuali adottate secondo schemi di auto o co-regolamentazione dai fornitori di piattaforme per la condivisione di video deve essere comunque valutata in ultima analisi dalle Autorità indipendenti competenti o agli organismi nazionali di regolamentazione.

Con riferimento al nuovo Codice europeo delle comunicazioni elettroniche sono state introdotte diverse misure che potenzialmente potrebbero impattare sulle piattaforme *online* il cui modello di *business* si basa sulla raccolta di dati. Ad esempio, nel Ritenuto 16 della Direttiva è stato rilevato che "i *servizi di comunicazione elettronica sono spesso forniti all'utente finale non solo in cambio di denaro, ma in misura sempre maggiore e in particolare in cambio della comunicazione di dati personali o di altri dati. Il concetto di remunerazione dovrebbe pertanto ricomprendere le situazioni in cui il fornitore di un servizio chiede all'utente finale dati personali ai sensi del regolamento (UE) 2016/679 o altri dati, e questi glieli trasmette consapevolmente, per via diretta o indiretta.*".

In base al riconoscimento della cessione dei dati (ad esempio, l'indirizzo IP, o altre informazioni generate automaticamente, come quelle raccolte e trasmesse da un *cookie*) come mezzo di pagamento, andranno a cadere nel perimetro di applicazione delle misure previste dal Nuovo Codice Europeo anche i cosiddetti "*servizi di comunicazione interpersonale*", ovvero quei servizi che consentono lo scambio interpersonale e interattivo di informazioni, come tutti i tipi di messaggi di posta elettronica, i servizi di messaggistica o le *chat* di gruppo, dietro la contropartita della cessione di dati da parte dell'utente finale. In particolare, per la prima volta ricadono nel perimetro regolamentare delle comunicazioni elettroniche anche i servizi di comunicazione interpersonale indipendenti da numero, ovvero quelle applicazioni che utilizzano il numero come identificativo (VoIP, messaggistica arricchita). Nel Considerando 18 del Nuovo Codice si stabilisce tuttavia che l'uso "indiretto" della numerazione non dovrebbe essere considerato equivalente all'uso di un numero per la connessione a numeri assegnati pubblicamente. Per questo motivo, i servizi di comunicazione interpersonale indipendenti dal numero dovrebbero essere soggetti a obblighi solo laddove interessi pubblici richiedano l'applicazione di obblighi normativi specifici a tutti i tipi di servizi di comunicazione interpersonale, indipendentemente dal fatto che utilizzino numeri per la fornitura del servizio. Nonostante il fatto che a questi servizi non venga attribuito lo stesso *status* dei servizi di comunicazione elettronica, il recente quadro normativo prevede tuttavia delle novità significative. La più importante è quella introdotta nel Considerando 149. In quest'ambito si riconosce uno dei principi-cardine della regolamentazione delle comunicazioni elettroniche, ovvero che sia la connettività da punto a punto e sia l'accesso ai servizi di emergenza, richiedono che gli utenti finali utilizzino servizi di comunicazione interpersonale *basati sul numero*. In questo contesto, la sempre maggiore diffusione di servizi di comunicazione interpersonale indipendenti dal numero potrebbe comportare un'*interoperabilità insufficiente tra servizi di comunicazione*, a danno degli utenti finali.

Per questo motivo, all'art. 61, comma 2, lettera c) del Nuovo Codice si stabilisce che, in casi giustificati, se la connettività da punto a punto tra gli utenti finali è compromessa a causa della mancanza di interoperabilità tra i servizi di comunicazione interpersonale, le Autorità possono imporre degli obblighi per i fornitori di servizi di comunicazione interpersonale indipendenti dal numero che abbiano un significativo livello di copertura e di diffusione tra gli utenti, al fine di rendere interoperabili i propri servizi. Tuttavia, tali obblighi possono essere imposti solo se proporzionati e se la Commissione, dopo aver consultato il BEREC e aver preso nella massima considerazione il suo parere, abbia riscontrato la presenza di una *notevole minaccia alla connettività da punto a punto tra*

utenti finali in tutta l'Unione o in almeno tre Stati membri e abbia adottato *misure di attuazione* che specificano le caratteristiche e la portata degli obblighi che possono essere imposti.

Con riferimento al Regolamento 2019/1150, questo promuove l'equità e la trasparenza per gli utenti commerciali dei servizi di intermediazione online (servizi *Platform to Business* o P2B) utilizzando un approccio e strumenti di azione tipici della regolamentazione *ex ante*. Nel Regolamento, sono infatti previsti specifici obblighi in capo ai "*fornitori di servizi di intermediazione online*" e ai "*fornitori di motori di ricerca online*", unitamente alla presenza di meccanismi di tutela a favore degli utenti commerciali e alla possibilità per gli Stati Membri di declinare le misure applicabili alle violazioni delle relative prescrizioni. Tale Regolamento è nato in risposta all'osservazione del fatto che molto spesso si registrano comportamenti scorretti da parte dei fornitori dei servizi di intermediazione *online* nei confronti di piccole e medie imprese. In particolare, le *malpractice* più significative riguardano i cambiamenti non motivati dei termini e delle condizioni oppure le chiusure improvvise dei siti. Il Regolamento, pertanto, si propone di garantire termini e condizioni eque e trasparenti, nonché effettive possibilità di ricorso per chiunque si interfacci con tali piattaforme, stabilendo una serie di obblighi in capo ad alcune figure imprenditoriali, definite come "*fornitori di servizi di intermediazione online*" (e.g.: Amazon, E-Bay, Netflix, Booking) e come "*fornitori di motori di ricerca online*" (e.g.: Google, Bing).

Con riferimento agli obblighi previsti dal Regolamento in capo ai fornitori di servizi di intermediazione *online*, i principali obblighi previsti riguardano la chiarezza nella redazione dei termini e delle condizioni, la comunicazione appropriata agli utenti di qualunque modifica di tali termini e condizioni, la previsione di meccanismi di comunicazione per limitazioni, sospensioni o cessazioni dei servizi, nonché la fissazione dei principali parametri che determinano il posizionamento di un certo prodotto/servizio, che debbono essere motivati. Per quanto riguarda i fornitori di motori di ricerca *online*, le misure più significative sono quelle relative all'obbligo di indicare i principali parametri più significativi per determinare il posizionamento di un certo servizio/prodotto sul motore di ricerca, specificando l'importanza relativa di tali parametri.

4. I *Big Data* nell'ecosistema digitale italiano: considerazioni del Garante per la protezione dei dati personali

4.1. Premessa

Al di là di quanto già anticipato nelle pagine precedenti con riguardo alla descrizione del fenomeno *Big Data*, considerazioni particolari devono essere svolte nei casi in cui le informazioni oggetto delle varie operazioni di trattamento riguardino "dati personali", locuzione da intendersi nella lata accezione accolta, tenuto conto delle fattispecie esaminate, non solo dalle autorità di protezione dei

dati personali[90], ma anche dagli orientamenti giurisprudenziali delle Corti superiori, sia della Corte di Giustizia[91] che della Corte di Cassazione[92].

Rispetto a tali ipotesi, infatti, e al di là di eventuali profili di possibile sovrapposizione con le tecniche di tutela proprie di altri ambiti (anzitutto quelli preordinati alla tutela del consumatore cui i trattamenti si riferiscano o sul quale gli stessi producano comunque un effetto), la disciplina di protezione dei dati personali, anzitutto quella ora contenuta nel RGPD, è un essenziale banco di prova per chi intenda esplorare le potenzialità offerte dai *Big Data* (anche in ragione del rigoroso quadro sanzionatorio che caratterizza il RGPD).

L'idea – che fa capolino in letteratura e che pure è talora serpeggiata nel corso delle audizioni – secondo cui l'ultimo "scossone" tecnologico derivante dai *Big Data* possa rappresentare la "testa d'ariete" per far capitolare le difese apprestate dagli ordinamenti europei rispetto alla tutela dei diritti fondamentali mediante l'ombrello offerto dalle discipline di protezione dei dati personali non è condivisa dal Garante. Al contrario, proprio la circostanza che la dimensione tecnologica acquisti con i *Big Data* una imponente capacità di spiegare i propri effetti (non tutti e non sempre benefici) sui singoli e sulla società nel suo complesso impone, nel nostro ordinamento costituzionale, non diversamente da quello dell'Unione europea, di preservare, oggi più che mai, le garanzie nel tempo acquisite a tutela dei diritti fondamentali e da ultimo ribadite nel RGPD.

L'approfondimento effettuato grazie all'Indagine svolta, con l'ausilio di quanti sono intervenuti e dei rappresentanti delle Autorità congiuntamente coinvolte, rappresenta un punto di avvio e non un punto di arrivo della riflessione in materia, consentendo una prima messa a fuoco degli aspetti più rilevanti

[90] Sulla (ampia) nozione di dato personale, si rinvia al sopra richiamato WP 136 del Gruppo art. 29, Parere 4/2007 sul concetto di dati personali adottato il 20 giugno 2007.

[91] Con riguardo, ad esempio, agli indirizzi IP, nella sentenza del 19 ottobre 2016 la Corte di Giustizia (II sez, causa C-582/14) ha stabilito che "*un indirizzo di protocollo Internet dinamico registrato da un fornitore di servizi di media online in occasione della consultazione, da parte di una persona, di un sito Internet che tale fornitore rende accessibile al pubblico costituisce, nei confronti di tale fornitore, un dato personale ai sensi di detta disposizione, qualora detto fornitore disponga di mezzi giuridici che gli consentano di far identificare la persona interessata grazie alle informazioni aggiuntive di cui il fornitore di accesso a Internet di detta persona dispone*". In modo ancora più rigoroso l'articolo 5, paragrafo 3, della direttiva 2002/58, dispone che gli Stati membri assicurano che l'archiviazione di informazioni oppure l'accesso a informazioni già archiviate nell'apparecchiatura terminale di un utente, "*senza qualificare tali informazioni né precisare che queste debbano essere dati personali*", siano consentiti unicamente a condizione che l'utente in questione abbia espresso preliminarmente il proprio consenso, dopo essere stato informato in modo chiaro e completo, a norma della direttiva 95/46, tra l'altro sugli scopi del trattamento (cfr. al riguardo la citata sentenza della Corte di Giustizia, 1° ottobre 2019, C-673/17, *Bundesverband der Verbraucherzentralen und Verbraucherverbände – Verbraucherzentrale Bundesverband eV c. Planet49 GmbH*, punti 66-71); sul trattamento di dati personali effettuato da Facebook riferito alle persone che visitano le *fanpage* presenti su *social network* grazie a marcatori («cookie» salvati da Facebook sul disco fisso del computer o su qualsiasi altro supporto dei visitatori della *fanpage*) contenenti ciascuno un codice utente unico che può essere associato ai dati di collegamento degli utenti registrati su Facebook, raccolto ed elaborato al momento dell'accesso alle *fanpage*. v. Corte di Giustizia, 5 giugno 2018, C-210/16, *Unabhängiges Landeszentrum für Datenschutz Schleswig-Holstein / Wirtschaftsakademie Schleswig-Holstein*.

[92] Accoglie una accezione ampi della nozione di dato personale anche Cass. civ., sez. II, 2 settembre 2015, n. 17440, che include nel novero dei dati personali anche le immagini raccolte mediante sistemi di videosorveglianza, in considerazione della possibilità di identificare i soggetti ripresi, a prescindere dalla notorietà degli stessi. In tempi più ravvicinati Cass. civ., Sez. II, 5 luglio 2018, n. 17665, ribadisce che la definizione di "dato personale" è un concetto molto ampio (e diverso da quello di "dato identificativo", che ne è una specie della categoria generale), in quanto ricomprensivo di qualsiasi informazione che consenta di identificare una persona fisica, tra cui il nome, il cognome e, nel caso considerato, l'indirizzo di posta elettronica.

al fine di individuare le modalità grazie alle quali non disperdere i benefici che un uso responsabile dei *Big Data* è in grado di offrire[93].

4.2. Gli interventi dei soggetti istituzionali

Del resto, un analogo processo è in corso da qualche anno presso larga parte delle autorità di protezione dei dati personali e delle istanze sovranazionali che pongono al centro della propria missione istituzionale la tutela dei diritti fondamentali.

Con la progressiva presa di conoscenza dei termini delle questioni sollevate dai *Big Data* – per quanto ancora rimangano in ombra gli effetti reali derivanti dai trattamenti in esame – si sono così moltiplicati gli approfondimenti e le raccomandazioni in materia. Ai documenti più risalenti, che si caratterizzano per maggiore sobrietà, se ne sono nel tempo succeduti altri che, in maniera più strutturata, individuano i profili critici connessi all'utilizzo dei *Big Data* e rendono avvertiti quanti ne facciano uso delle misure da adottare. Alla luce di tali documenti e linee guida, elaborati in tempi recenti – taluni dei quali, senza pretesa di esaustività, si indicano di seguito –, è già possibile sviluppare una prima strategia integrata (e, va aggiunto, sufficientemente condivisa) orientata a consentire l'utilizzo virtuoso dei *Big Data* nel rispetto del diritto alla protezione dei dati personali e di ritrarre prime indicazioni, (inevitabilmente) suscettibili di affinamento alla luce dell'applicazione concreta, in prospettiva anche da parte del Comitato europeo per la protezione dei dati. Si tratta, in particolare, dai seguenti documenti (taluni dei quali peraltro già richiamati):

- Council of Europe's Recommendation CM/Rec (2010)13 *on the protection of individuals with regard to automatic processing of personal data in the context of profiling*;
- International Working Group on Data Protection in telecommunications (IWGDPT), *Working paper on Big Data and Privacy. Privacy principles under pressure in the age of Big Data analytics*, 5-6 May 2014;
- Dichiarazione del 16 settembre 2014 del Gruppo di lavoro sulla protezione dei dati "Articolo 29" relativa all'impatto dello sviluppo dei Big Data sulla protezione delle persone rispetto al trattamento automatizzato dei loro dati personali nell'UE;
- 36a Conferenza internazionale delle autorità di protezione dei dati, Mauritius 2014, *Resolution Big Data;*
- EDPS: Preliminary Opinion on Privacy and Competitiveness in the Age of Big Data, 2014; Opinion 4/2015 *Towards a new digital ethics. Data, dignity and technology*; Opinion 7/2015 *Meeting the challenges of big data. A call for transparency, user control, data protection by design and accountability*; Opinion 8/2016 *on coherent enforcement of fundamental rights in the age of big data*; Opinion 3/2018 *on online manipulation and personal data*;
- Council of Europe, Consultative Committee of the Convention for the Protection of Individuals with Regard to Automatic Processing of Personal Data (T-PD), *Guidelines on the protection of*

[93] Nella stessa direzione, con riguardo alle ipotesi (allo studio) di condivisione dei dati tra le imprese e il settore pubblico, nella Comunicazione della Commissione al Parlamento europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle regioni, *Verso uno spazio comune europeo dei dati*, Bruxelles, 25.4.2018, COM(2018) 232 final, si ribadisce che "*qualsiasi azione intrapresa a questo riguardo deve essere pienamente conforme alla legislazione in materia di protezione dei dati personali*" (p. 13).

individuals with regard to the processing of personal data in a world of Big Data, Strasbourg, 23 January 2017 T-PD(2017)01;

- Risoluzione del Parlamento europeo del 14 marzo 2017 sulle *implicazioni dei Big Data per i diritti fondamentali: privacy, protezione dei dati, non discriminazione, sicurezza e attività di contrasto* (2016/2225(INI)).

Nel solco di questi interventi intende porsi anche il Garante. E del resto anche soggetti istituzionali operanti a livello globale hanno avvertito da tempo e segnalato le gravi implicazioni di natura non solo individuale, ma collettiva che uno sviluppo non "sorvegliato" dei *Big data* può determinare, ben potendo le stesse riguardare la struttura democratica delle società occidentali (come peraltro in tempi recenti la vicenda Cambridge Analytica ha disvelato): "*The wealth of individual-level information that Google, Facebook, and a few mobile phone and credit card companies would jointly hold if they ever were to pool their information is in itself concerning. Because privacy is a pillar of democracy, we must remain alert to the possibility that it might be compromised by the rise of new technologies, and put in place all necessary safeguards. […] Any initiative in the field ought to fully recognise the salience of the privacy issues and the importance of handling data in ways that ensure that privacy is not compromised*"[94].

L'Indagine svolta ha fatto anche emergere la possibilità di fruttuose interazioni tra autorità amministrative indipendenti[95]: pur restando (necessariamente) distinte le attribuzioni rimesse per legge a ciascuna di esse[96], come pure le finalità ultime perseguite dalle discipline che governano i rispettivi ambiti di azione, un'attiva collaborazione tra autorità amministrative indipendenti di volta in volta competenti (che ben potrebbero non esaurirsi in quelle che hanno partecipato alla presente iniziativa) può contribuire ad un superamento dei limiti connaturati alla prospettiva "in silos" che abitualmente ne caratterizza l'azione. Da questo punto di vista, la materia della protezione dei dati personali si pone, per la trasversalità che la caratterizza, come punto di incrocio necessario rispetto a tutti gli ambiti regolatori interessati dal fenomeno *Big Data* (che non si esauriscono nei trattamenti

[94] Cfr. UN Global Pulse, *Big Data for Development: Challenges & Opportunities*, May 2012, p. 24.

[95] Già con riguardo alle interrelazioni tra la disciplina di protezione dei dati e quella a tutela della concorrenza v. per tutti G. Pitruzzella, *Big data, competition and privacy: a look from the antitrust perspective*, in *Concorrenza e mercato*, 2016, 15; Id., *Big Data and antitrust enforcement*, in *Riv. It. Antitrust*, 2017, fasc. 1, 10; Colangelo G., Maggiolino M., *Big Data, data protection and antitrust in the wake of the "Bundeskartellamt" case against Facebook*, in *Riv. It. Antitrust*, 2017, fasc. 1, 9.

[96] In questa prospettiva, ad esempio, il Tar Lazio (sez. I) con sentenza 7 maggio 2018, n. 5043, in relazione ad una vicenda connessa a manifestazioni promozionali caratterizzate dalla promessa di ulteriori prodotti e/o rimborsi sul prezzo beni pubblicizzati, ha parzialmente accolto il ricorso e per l'effetto annullato un provvedimento dell'Autorità garante della concorrenza e del mercato nella parte relativa ad una presunta pratica commerciale scorretta posta in essere da una società con riguardo all'acquisizione, per finalità di marketing, del consenso all'utilizzo dei dati personali dei consumatori necessario per richiedere i premi/vantaggi promessi. Il TAR ha affermato che «*l'eventuale illegittima raccolta dei dati presenti nella piattaforma o la loro cattiva gestione da parte del professionista costituisce una possibile violazione dei principi in materia di corretto trattamento dei dati personali, il cui accertamento non è di competenza dell'Agcm, trattandosi di una prerogativa rimessa ai sensi del d.lgs. 30 giugno 2003, n. 196 al Garante per la protezione dei dati personali. L'Agcm è, invece, chiamata a verificare se la richiesta dei dati sia avvenuta secondo modalità tali da condizionare la libertà di scelta del consumatore, interessato solo all'ottenimento del premio relativo a una campagna promozionale, obbligandolo anche a partecipare al programma di fidelizzazione della clientela*» (v. p. 10-11). Più in generale, sui possibili effetti di rimbalzo (virtuoso) tra decisioni di Autorità indipendenti adottate nei rispettivi ambiti di competenza di segnala il provv. Garante 22 maggio 2018, n. 363, doc. web n. 8995274, adottato a carico di una società cui ha fatto seguito la Delibera n. 391/18/CONS dell'Autorità per le Garanzie nelle Comunicazioni, in virtù della quale si è determinata, nei confronti della medesima società, la decadenza dell'accreditamento di un motore di calcolo di cui alla delibera n. 22/10/CONS a seguito delle violazioni in materia di protezione dei dati personali accertate dal Garante.

effettuati da soli soggetti privati o in quelli che hanno un impatto sui soli consumatori, atteso il possibile utilizzo di tali tecniche anche in ambito pubblico) e, dal punto di vista operativo, come partner "naturale" sia nella prospettiva di ulteriori approfondimenti, sia nell'assolvimento dei compiti di *enforcement*.

4.3. Oltre la pura descrizione del fenomeno

Anche dall'Indagine svolta (i cui termini si sono sintetizzati nelle pagine che precedono), come pure dalla ricognizione della letteratura più recente in materia (nella quale però è marcato l'*imprinting* nord-americano, ordinamento nel quale il fenomeno si è fatto strada veicolato dalle *Big tech* in assenza, è bene tenerlo presente approcciando la materia in esame, di un quadro normativo in materia di protezione dei dati personali comparabile con le garanzie offerte da quello europeo), emerge la necessità di una più precisa messa a fuoco – anzitutto da parte degli *stakeholder* istituzionali – delle interrelazioni tra il fenomeno dei *Big Data* e la disciplina di protezione dei dati personali e, più precisamente, dei singoli istituti e rimedi che la compongono.

Tale operazione va di pari passo con la necessità di calare le ipotesi astrattamente riconducibili nella categoria (dalla valenza puramente descrittiva dei) *Big Data* in singole concrete applicazioni, suscettibili (in ragione delle peculiarità che le possono connotare) di diverso apprezzamento sul piano degli effetti giuridici[97]: non solo le attività di controllo che il Garante potrà effettuare in questi ambiti, ma anche i pareri resi a seguito di consultazione preventiva (art. 36 RGPD) potranno costituire utili occasioni per verificare in concreto oltre alle potenzialità anche i rischi connessi all'utilizzo dei *Big Data*. Rischi già paventati sui diritti degli interessati (anche se sovente evocati in letteratura in termini assai generici) nonché per la *privacy* individuale (la formula è qui volutamente utilizzata nella sua massima ampiezza); rischi che talora prendono forma con riguardo all'adozione di decisioni, che riposano su trattamenti effettuati con tecniche *Big Data*, di natura discriminatoria rispetto a singoli o classi di interessati[98].

E se pure le discipline di protezione dei dati personali (prima la direttiva 95/46 ed ora il RGPD) di rado formano oggetto di puntuale declinazione rispetto alla materia in esame, rimanendo sullo sfondo delle trattazioni dedicate al tema[99], pur si vanno affacciando riflessioni secondo le quali, in prima approssimazione, le attività legate all'utilizzo dei *Big Data* possono evidenziare chiari profili di contrasto con aspetti fondamentali della disciplina di protezione dei dati, anzitutto con riferimento ai principi di liceità e correttezza nel trattamento, aspetto quest'ultimo che rinvia ad una effettiva (e compiuta) consapevolezza degli interessati (e correlativa trasparenza dei titolari del trattamento) circa le operazioni di trattamento dei dati personali connesse all'utilizzo dei dati personali che li potrà

[97] Si pensi, ad esempio, all'accertamento della sussistenza del requisito del legittimo interesse di cui all'art. 6, par. 1, lett. *f*), del RGPD, quale condizione di legittimità del trattamento che, se non può essere aprioristicamente esclusa rispetto a singole manifestazioni del fenomeno in esame, neanche può ritenersi comunque ricorrente rispetto ad ogni trattamento effettuato con le modalità prese qui in considerazione: tale valutazione, da effettuarsi in prima istanza da parte del titolare del trattamento, non può infatti non tenere conto di tutti gli elementi che in concreto caratterizzano i singoli ambiti applicativi delle tecniche di *Big Data*. Tale base giuridica del trattamento responsabilizza peraltro il titolare del trattamento, non solo perché deve dimostrare la prevalenza del proprio interesse rispetto ai diritti degli interessati, ma anche perché, in base al nuovo quadro normativo introdotto con il RGPD, tale circostanza deve essere chiarita nell'informativa resa agli interessati (cfr. artt. 13, par. 1, lett. *d*) e 14, par. 2, lett. *b*), del RGPD.

[98] V. già al riguardo FTC Report, *Big Data. A Tool for Inclusion or Exclusion? Understanding the Issues*, January 2016.

[99] Nel corso delle audizioni ha evidenziato questa carenza anche J. Cannataci il 15 gennaio 2018. V. anche la recente ricognizione di G.M. Ruotolo, *I dati non personali: l'emersione dei big data nel diritto dell'Unione europea*, in *Studi integr. eur.*, XIII (2018), 97.

riguardare; alla violazione del principio di finalità; alla corretta individuazione della base giuridica posta a fondamento di tali operazioni di trattamento, anzitutto con riguardo al consenso degli interessati[100]. Anche i principi di minimizzazione, di limitazione della finalità e di conservazione dei dati per il solo tempo indispensabile alla realizzazione del trattamento mal si attagliano a raccolte massive di dati, in ipotesi acquisiti, magari non per esigenze attuali ma in vista di future e solo ipotetiche necessità, per essere quindi riutilizzati per fini ulteriori non sempre compatibili con quelli originari; l'utilizzo di algoritmi complessi nel processo di *Big Data analytics* può comportare risultati inattesi, che potrebbero essere lesivi di interessi individuali e porsi in violazione del principio di correttezza. Quanto all'ulteriore profilo di dettaglio, merita ricordare che anche il Parlamento europeo nella risoluzione del 14 marzo 2017, ha messo in evidenza il rischio che anche la distinzione (propria delle normative di protezione dei dati personali) tra dati sensibili e non venga a sfumare nel mondo dei *Big Data*, potendo i primi essere estratti combinando tra loro dati comuni.

La stessa nozione di "dato anonimo" (quale limite esterno delle garanzie accordate dalla disciplina di protezione dei dati) richiede un vaglio attento a seconda delle caratteristiche che, in concreto, ciascun trattamento effettuato secondo la metodologia *Big Data* potrà in concreto presentare.

Considerazioni particolari dovrebbero poi essere svolte rispetto all'utilizzo di tecniche *Big Data* con *dataset* di provenienza pubblica ovvero effettuati dalle pubbliche amministrazioni, profilo rientrato però nel compasso dell'Indagine congiunta (cfr. comunque, per prime considerazioni, il successivo par. 4.13).

4.4. Le implicazioni etiche

Non stupisce allora – ed anzi costituisce un positivo invito alla riflessione e all'approfondimento critico (anche in ordine alla congruità degli strumenti posti a disposizione delle autorità di controllo per affrontare gli inediti rischi sollevati dai *Big Data*) − che nella segnalata cornice normativa, da taluno avvertita come inadeguata rispetto al fenomeno preso in esame[101], da un lato, con apprezzabile realismo, si formulino inviti a monitorare il fenomeno in ragione della sua (relativa) novità[102]; dall'altro, si sia rimarcata la necessità di un approfondimento delle implicazioni etiche connesse all'introduzione dei *Big Data*[103] che, ove non adeguatamente "governati", possono alimentare (più di

[100] Con riguardo a fenomeni prossimi ai *Big data*, tali aspetti sono stati toccati in numerosi documenti adottati dal Gruppo art. 29, la cui attualità non è venuta meno con il RGPD: v., in particolare, WP 251 rev.01, *Linee guida sul processo decisionale automatizzato relativo alle persone fisiche e sulla profilazione ai fini del regolamento 2016/679* adottate il 3 ottobre 2017. Versione emendata e adottata in data 6 febbraio 2018; in precedenza, cfr. già Parere n. 2/2006 sugli aspetti di tutela della vita privata inerenti ai servizi di screening dei messaggi di posta elettronica adottata il 21 febbraio 2006; WP 171 Parere 2/2010 sulla pubblicità comportamentale online, adottato il 22 giugno 2010; Parere n. 9/2014 sull'applicazione della Direttiva 2002/58/EC al *device fingerprinting* adottata il 25 novembre 2014.

[101] V. ad es. G. De Minico, *Does the European Commission's decision on Google open new scenarios for the Legislator?*, in *Osservatorio costituzionale*, 2017, fasc. 3, 6.

[102] Cfr. Joint Committee of the European Supervisory Authorities, *Joint Committee Final Report on Big Data*, JC/2018/04, 15 March 2018.

[103] Cfr. G. Buttarelli, *Le sfide dei big data tra evoluzione tecnologica, etica e interessi collettivi*, in *Gnosis*, 2017, 31, il quale si domanda se, entro questa cornice in trasformazione, non vi sia spazio per:

- "la codificazione in chiave evolutiva di princìpi orizzontali o settoriali;
- l'individuazione di soggetti preposti alla loro enucleazione (i gestori stessi delle informazioni, oppure organi pubblici, autorità preposte o entità 'paritetiche');
- l'attribuzione agli stessi di un qualche valore cogente oppure basato sull'autodisciplina;
- la comprensione di se e come, anche per gli organismi pubblici e di *enforcement*, sia utile fare riferimento a tali valori".

quanto i trattamenti in essere già non contribuiscano a fare) un regime della classificazione e della sorveglianza. È a questo proposito necessario guardare agli *effetti* sui singoli derivanti dall'utilizzo dei *Big Data* – specie ove, a seguito di attività di profilazione (cfr. Considerando 71 e art. 4, n. 4, RGPD), possano condurre all'adozione di misure discriminatorie[104] –, ma anche alle *aspettative* individuali in relazione al potenziale utilizzo di *dataset* contenenti dati personali che li riguardano, potendo essere minata, con usi inattesi delle informazioni, la fiducia riposta nel titolare del trattamento (che, va ribadito, è mero "custode" e non *dominus* dei dati personali raccolti per finalità che la legge vuole determinate e rese note all'interessato e quindi non arbitrariamente modificabili).

Già solo sul piano etico, la trasparenza dei processi che si avvalgono di tecniche *Big Data* – valore al quale dà corpo la disciplina relativa agli obblighi informativi contenuta nel RGPD – rappresenta un fattore chiave il cui perseguimento è irrinunciabile, sia nel settore privato[105] che in quello pubblico.

4.5. *Big Data*, principio di qualità dei dati (e dei processi) e profilazione

Ma pure l'affidabilità e la qualità dei processi di *Big Data analysis* rappresenta un pre-requisito per tali trattamenti e per la successiva applicazione su larga scala delle "conoscenze" acquisite. In proposito va ricordato che, con specifico riguardo all'attività di profilazione, il Considerando 71 del RGPD evidenzia che "tenendo in considerazione le circostanze e il contesto specifici in cui i dati personali sono trattati, è opportuno che il titolare del trattamento utilizzi *procedure matematiche o statistiche appropriate per la profilazione*, metta in atto misure tecniche e organizzative adeguate al fine di garantire, in particolare, che siano rettificati i fattori che comportano inesattezze dei dati e sia minimizzato il rischio di errori e al fine di garantire la sicurezza dei dati personali secondo una modalità che tenga conto dei potenziali rischi esistenti per gli interessi e i diritti dell'interessato e impedisca, tra l'altro, effetti discriminatori nei confronti di persone fisiche sulla base della razza o dell'origine etnica, delle opinioni politiche, della religione o delle convinzioni personali, dell'appartenenza sindacale, dello status genetico, dello stato di salute o dell'orientamento sessuale, ovvero un trattamento che comporti misure aventi tali effetti".

Il principio di qualità gioca un ruolo decisivo in relazione all'utilizzo di tecniche di *Big Data analytics* per la profilazione degli interessati. Il Garante, già prima dell'entrata in vigore del RGPD, era intervenuto in materia di profilazione, ribadendo l'importanza dei principi fondamentali alla base di ogni trattamento ed evidenziando i particolari rischi di una profilazione effettuata tramite Internet [106].

Con l'entrata in vigore del RGPD la "profilazione" è definita come "*qualsiasi forma di trattamento automatizzato di dati personali consistente nell'utilizzo di tali dati personali per valutare determinati aspetti personali relativi a una persona fisica, in particolare per analizzare o prevedere aspetti*

Sulla dimensione etica delle sfide poste dai *Big Data*, profilo sollevato tra i primi dall'Edps con proprie iniziative, v. pure M. Kelly, P. Twomey, *I big data e le sfide etiche*, in *La civiltà cattolica*, 2018, 40.

[104] Si incentra su questi aspetti, ad esempio, il rapporto curato dalla Federal Trade Commission, *Big data: a tool for inclusion or exclusion*, 2016.

[105] Criticamente, in relazione alle modalità che risulterebbero essere state seguite nell'interesse di Google in vista del miglioramento degli algoritmi di *facial recognition*, v. le notizie di stampa pubblicate su https://www.theguardian.com/technology/2019/oct/03/google-data-harvesting-facial-recognition-people-of-color; v. pure https://www.nytimes.com/2019/10/04/technology/google-facial-recognition-atlanta-homeless.html.

[106] Cfr. Linee guida in materia di trattamento di dati personali per profilazione on line del 19 marzo 2015, pubblicato sulla Gazzetta Ufficiale n. 103 del 6 maggio 2015, doc. web 3881513 e, in particolare: il rilascio di una idonea informativa, l'acquisizione di un consenso specifico da parte degli interessati, la sussistenza del diritto di opposizione e il principio di finalità.

riguardanti il rendimento professionale, la situazione economica, la salute, le preferenze personali, gli interessi, l'affidabilità, il comportamento, l'ubicazione o gli spostamenti di detta persona fisica", anche attraverso un tracciamento delle persone fisiche su internet (art. 4 e considerando 24 e 30).

Come noto, tra le caratteristiche principali, che determinano il valore aggiunto dei *Big Data* rispetto alle tradizionali forme di profilazione, vi è la possibilità di utilizzare dati raccolti in modo automatizzato attraverso l'osservazione dell'interessato – si pensi ai dati relativi al comportamento *online* attraverso l'impiego dei *cookies*[107] e altri marcatori relativi ad attività *online* ovvero ai dati generati dai dispositivi della *Internet of Things* – e dati inferiti (o derivati), cioè dati che vengono dedotti da altri dati utilizzando tecniche di analisi volte ad individuare correlazioni statistico-probabilistiche ricorrenti un dato *dataset*.

Pertanto, le tecniche automatizzate di *Big Data analytics* basate su dati personali possono essere impiegate per identificare *trend* comportamentali degli interessati e per estrarre conoscenza predittiva, allo scopo di orientare decisioni in riferimento a persone o gruppi[108]. Tale processo determina implicazioni in riferimento al rispetto dei diritti fondamentali, tra i quali il diritto alla vita privata, alla protezione dei dati e alla loro sicurezza, la libertà di espressione e di non discriminazione[109]. In tale ottica le *Linee guida sul processo decisionale automatizzato relativo alle persone fisiche e sulla profilazione* prendono le mosse dall'assunto che l'evoluzione tecnologica, tra cui l'analisi dei *Big Data*, ha semplificato la possibilità di creare profili degli interessati e decisioni automatizzate sulla base di risultanze degli algoritmi di profilazione: "*i progressi tecnologici e le capacità in materia di analisi dei megadati (big data), intelligenza artificiale e apprendimento*

[107] La materia ha formato oggetto di un rilevante (anche per la materia dei *Big Data*, non di rado frutto nel mondo *online* di un processo di accumulazione generato proprio dai *cookies* di terze parti) e recente intervento della Corte di giustizia dell'Unione europea, con sentenza del 1° ottobre 2019, causa C-673/17 *Bundesverband der Verbraucherzentralen und Verbraucherverbände − Verbraucherzentrale Bundesverband eV / Planet49 GmbH*. A seguito di ricorso in via pregiudiziale del *Bundesgerichtshof* tedesco. La fattispecie considerata riguarda l'utilizzo, da parte di una società tedesca, nell'ambito di giochi a premi in linea, di una casella di spunta preselezionata mediante la quale gli utenti di Internet che intendono partecipare al gioco esprimono il loro accordo all'installazione di cookie diretti a raccogliere informazioni a fini pubblicitari. In relazione ad essa, la Corte ha dichiarato che il consenso che l'utente di un sito Internet deve prestare ai fini dell'installazione e della consultazione di cookie sulla sua apparecchiatura terminale non è validamente manifestato mediante una casella di spunta preselezionata che l'utente deve deselezionare al fine di negare il proprio consenso. La Corte rimarca in particolare il requisito della necessaria specificità del consenso da parte dell'utente il quale "*deve riferirsi precisamente al trattamento dei dati interessati e non può essere desunta da una manifestazione della volontà avente un oggetto distinto*" (punto 58). La Corte sottolinea che il consenso deve essere specifico, cosicché il fatto che un utente attivi il pulsante di partecipazione ad un gioco a premi non è sufficiente per ritenere che l'utente abbia validamente espresso il proprio consenso all'installazione di cookie (punto 59). Tale assunto risulterebbe ulteriormente confermato dal RGPD (richiamando al riguardo il considerando 32). Secondo la Corte, infine, il periodo di attività dei cookie, nonché la possibilità o meno per i terzi di avere accesso a tali *cookie* rientrano tra le informazioni che il fornitore di servizi deve comunicare all'utente, precisando che le informazioni contenute nell'art. 10 della direttiva 95/46, in ragione del tenore letterale della disposizione, "*non sono elencate in modo esaustivo*". Di qui l'assunto che "*l'informazione sul periodo di attività dei cookie deve essere considerata rispondente al requisito del trattamento leale dei dati previsto dal suddetto articolo, in quanto, in una situazione come quella di cui trattasi nel procedimento principale, un lungo periodo di attività, o addirittura un periodo illimitato, implica la raccolta di numerose informazioni sulle abitudini di navigazione e sulla frequenza delle eventuali visite dell'utente ai siti dei partner pubblicitari dell'organizzatore del gioco a premi*" (punto 78), con richiamo al riguardo dell'art. 13, par. 2, lett. a), RGPD. V. in merito la posizione enunciata il 14 novembre 2019 dall'Hamburg Commissioner for Data Protection and Freedom of Information, *Google Analytics and similar services can only be used with consent*.

[108] Consiglio d'Europa, Comitato consultivo della Convenzione n. 108, *Guidelines on the protection of individuals with regard to the processing of personal data in a world of Big Data*, 23 gennaio 2017, p. 2.

[109] Risoluzione del Parlamento europeo del 14 marzo 2017 sulle implicazioni dei Big Data per i diritti fondamentali: privacy, protezione dei dati, non discriminazione, sicurezza e attività di contrasto (2016/2225(INI)).

automatico hanno reso più facile la creazione di profili e l'adozione di decisioni automatizzate, con potenziali ripercussioni significative sui diritti e sulle libertà delle persone fisiche"[110].

Il RGPD, anche se non si occupa direttamente di *Big Data*, prevede disposizioni applicabili a tali ipotesi, volte a fronteggiare i potenziali rischi derivanti dalla profilazione e dal processo decisionale automatizzato e a tutelare i diritti fondamentali degli interessati, ponendo limitazioni nei casi in cui i *Big Data* possano avere un impatto significativo sugli individui.

Ai sensi del RGPD, anche la profilazione e i processi decisionali automatizzati devono essere svolti nel rispetto dei principi generali di liceità, correttezza e trasparenza, finalità, minimizzazione, esattezza, limitazione della conservazione e in presenza di una base giuridica per il trattamento (artt. 5 e 6).

Inoltre la profilazione, al pari di ogni altro trattamento, prevede specifici diritti in capo agli interessati, tra i quali il diritto di opposizione e il diritto di "*non essere sottoposto a una decisione basata unicamente sul trattamento automatizzato, compresa la profilazione, che produca effetti giuridici che lo riguardano o che incida in modo analogo significativamente sulla sua persona*"[111] (art. 21 e 22).

4.6. Per un approccio *win-win*

A fronte delle rilevate riserve o criticità (che pure hanno fatto capolino nel corso dell'Indagine[112]), non può tuttavia corrispondere un arretramento nel sistema delle tutele rispetto ai diritti presidiati dalle Autorità; al contrario, il Garante ritiene necessario – se del caso, grazie a forme rafforzate di cooperazione ed anche percorrendo, ove possibile e fruttuoso, la via dei codici di condotta – operare in senso opposto. Con ragione, a questo proposito, si è affermato che "*lo sviluppo dei Big Data può avere ripercussioni su libertà e diritti fondamentali delle persone se non accompagnato da garanzie e da principi etici che li rendano compatibili con i valori delle società democratiche. […] In generale, una Big Data strategy deve essere predisposta nella piena consapevolezza dei limiti, fissati anche dalle norme in materia di protezione dei dati personali, alle analisi conducibili e ai risultati ottenibili, per evitare di ledere diritti e libertà fondamentali delle persone*"[113]. In questa prospettiva pare potersi (opportunamente) inscrivere, ad esempio, il principio n. 8 della "Carta dei principi tecnologici del *procurement*", che mira a definire i criteri minimi per lo sviluppo di servizi digitali della Pubblica Amministrazione e secondo il quale (pur rimanendo ad un livello assai elevato di astrazione) è necessario "*assicurarsi che i diritti dei cittadini siano protetti integrando la privacy come parte*

[110] Cfr. WP 251 rev.01, Linee guida sul processo decisionale automatizzato relativo alle persone fisiche e sulla profilazione ai fini del regolamento 2016/679, adottate il 3 ottobre 2017. Versione emendata e adottata in data 6 febbraio 2018.

[111] L'art. 22, par. 2 introduce tre eccezioni al divieto generale di utilizzo del processo decisionale basato unicamente sul trattamento automatizzato, compresa la profilazione, quando: a) sia necessaria per la conclusione o l'esecuzione di un contratto tra l'interessato e un titolare del trattamento; b) sia autorizzata dal diritto dell'Unione o dello Stato membro cui è soggetto il titolare del trattamento, che precisa altresì misure adeguate a tutela dei diritti, delle libertà e dei legittimi interessi dell'interessato; c) si basi sul consenso esplicito dell'interessato.

[112] V. le riflessioni svolte dal Prof. A. Mantelero nel Testo scritto depositato sulla base delle risultanze dell'audizione del 21 novembre 2017, incentrate sui principi di protezione dei dati personali, focalizzatesi, in particolare rispetto al processo di adozione della Proposta di Regolamento *e-privacy*, destinato a modificare il panorama normativo derivante dalla direttiva 2002/58; su tale processo v. anche le osservazioni critiche formulate nel corso dell'indagine da parte degli operatori di telecomunicazione i quali hanno evidenziato la necessità di equiparazione con gli OTT in relazione all'offerta di servizi comparabili.

[113] In merito v. C. Comella, *Origine dei "Big Data"*, in *Gnosis*, 2017, 2, 130.

essenziale del sistema. Inserire nel capitolato l'obbligo di rispettare le prescrizioni della normativa italiana ed europea sulla protezione dei dati personali (RGPD)"[114].

Ai rischi generalmente paventati rispetto al fenomeno *Big Data* si deve allora rispondere con una rinforzata strategia di protezione dei dati personali, non però nella logica dell'aut-aut (vale a dire della contrapposizione), ma in quella dell'et-et (o se si vuole, del *win-win*)[115], anzitutto con l'adozione di adeguate misure tecnologiche ed organizzative − che potranno di regola essere individuate all'esito di una seria valutazione di impatto effettuata ai sensi dell'art. 35 del RGPD e, ove necessario, a seguito della consultazione dell'autorità di controllo ai sensi del successivo art. 36 − improntata ai principi (ora esplicitati nel Regolamento europeo) della *privacy by design* e *by default.*

4.7. L'opacità dei trattamenti con tecniche *Big Data* e il principio di trasparenza proprio delle discipline di protezione dei dati

Se generalizzata è la denuncia circa la (tendenziale) opacità dei trattamenti in esame[116] riconoscendosi anche in questa materia i limiti delle tradizionali *privacy policy*, è bene ricordare, denunciati per primi proprio dalle autorità di protezione dei dati personali –, è proprio a tale opacità che dovrà rispondersi valorizzando (ed inverando) anzitutto il principio di *accountability* – che si ritiene caratterizzare il RGPD –, atteso che l'assenza di informazioni con riguardo alle caratteristiche principali dei trattamenti di dati personali contrasta con uno degli assi portanti delle discipline di protezione dei dati: quello della trasparenza[117].

Peraltro, per trattare ulteriormente i dati personali per una finalità diversa – comunque non incompatibile con quella per cui essi sono stati raccolti (e abbiamo visto che frequentemente ciò può accadere rispetto ai *Big Data*) – il RGPD prevede un rigido quadro di garanzie (artt. 6, par. 4 e 23 del RGPD) in base alle quali, laddove non vi sia il consenso dell'interessato (in ambito privato) o un atto legislativo che costituisca una misura necessaria e proporzionata di una "società democratica" per la salvaguardia di alcuni rilevanti interessi pubblici (in ambito pubblico), il titolare del trattamento deve rispettare una serie di precise garanzie. In particolare, nella prospettiva della piena informazione sono espressamente orientati gli artt. 13, par. 3 e 14, par. 4, del RGPD, secondo i quali il titolare del trattamento, qualora intenda trattare ulteriormente i dati personali per una finalità diversa da quella per cui essi sono stati raccolti, prima di tale ulteriore trattamento fornisce all'interessato le informazioni in merito a tale diversa finalità e ogni ulteriore informazione pertinente.

E del resto (quantomeno) l'informazione – da rendersi "in forma concisa, trasparente, intelligibile e facilmente accessibile, con un linguaggio semplice e chiaro, in particolare nel caso di informazioni destinate specificamente ai minori" (art. 12, par. 1, del RGPD) – è l'elemento chiave per rispettare la dignità dell'interessato (oltre che gli interessi dei consumatori) in relazione all'utilizzo dei propri dati:

---

[114] Cfr. https://medium.com/team-per-la-trasformazione-digitale/carta-dei-principi-tecnologici-del-procurementpubblica-amministrazione-innovazione-dbdfcab2745.

[115] In questo senso v. il rapporto curato da ENISA, *Privacy by design in big data*, 2015

[116] Anche per F. Bernabè nell'*Intervento* al Convegno *Big Data e Privacy. La nuova geografia dei poteri*, cit., p. 23, "*Il tema dei Big Data pone soprattutto problemi di regolazione nell'accesso, nella elaborazione e nell'utilizzo dei dati. Una regolazione non finalizzata a limitare in qualche modo le potenzialità che ne derivano ma soprattutto a garantirne la trasparenza*".

[117] V. al riguardo i contributi raccolti negli Atti del Convegno organizzato dal Garante, *Big Data e Privacy. La nuova geografia dei poteri*, 30 gennaio 2017, Roma, 2017, dove si rinvengono sollecitazioni a favore dell'introduzione di garanzie di trasparenza dei processi, a contrastare la crescente difficoltà a mantenere un effettivo controllo sui dati, sia in ragione dell'opacità delle modalità di raccolta, sia in ragione dei criteri di selezione e di analisi dei dati.

la qualità effettiva della stessa e le modalità procedimentali utilizzate rappresenteranno, specie nel contesto dei *Big Data*, i termini di riferimento per misurare l'osservanza ai principi di correttezza e trasparenza (art. 5, par. 1, lett. *a*), del RGPD). Del resto la disciplina di protezione dei dati non richiede, neanche in relazione ai *Big Data*, che vengano rappresentati all'interessato le modalità di trattamento dei dati (vale a dire i dettagli tecnici mediante i quali le operazioni di trattamento sono effettuate), quanto piuttosto le finalità che sono in concreto perseguite.

Nel peculiare contesto qui considerato merita sottolineare tra le informazioni da rendersi all'interessato quelle concernenti gli eventuali "*legittimi interessi perseguiti dal titolare del trattamento o da terzi*" qualora il trattamento si basi sull'articolo 6, paragrafo 1, lettera *f*), del RGPD (fattispecie più volte evocata nel corso delle audizioni). I legittimi interessi, affinché il trattamento possa ritenersi lecito sulla base di tale presupposto giuridico, non devono peraltro prevalere sugli interessi, i diritti e le libertà delle persone che richiedono la protezione dei dati personali, sulla base di un *test* comparativo in cui l'impatto sugli interessati va rigorosamente valutato[118].

Del pari, aspetto esso pure rilevante con riguardo ai *Big Data*, all'interessato va altresì rappresentata l'eventuale "*esistenza di un processo decisionale automatizzato, compresa la profilazione di cui all'articolo 22, paragrafi 1 e 4, e, almeno in tali casi,* [vanno fornite] *informazioni significative sulla logica utilizzata, nonché l'importanza e le conseguenze previste di tale trattamento per l'interessato*".

Né tali obblighi informativi rappresentano inutili ed onerosi adempimenti burocratici, adducendo la circostanza (di fatto) che i più non scorrono le cd. *privacy policy*. Bisogna ribaltare l'idea che gli obblighi informativi rappresentino un *red carpet*: non lo sono. Da un lato perché informative insoddisfacenti possono per ciò solo integrare una violazione della disciplina di protezione dei dati personali (e spesso sono sintomo di trattamenti che, proprio in quanto non esplicitati, possono presentare profili di criticità); dall'altro, perché le *privacy policy* sono lo strumento principe destinato a "fotografare" il trattamento che verrà posto in essere e, sulla scorta delle informazioni fornite, l'interessato (o il consumatore, se si vuole) non è lasciato solo e impotente: ha lo strumento del reclamo alle autorità di controllo che, nell'esercizio della propria missione istituzionale, potranno supplire alla oggettiva situazione di subalternità dei singoli. Esse, proprio muovendo dall'informativa resa agli interessati, potranno infatti trarre spunto ed elementi utili per accertare la complessiva liceità del trattamento.

Le informazioni rese agli interessati, del resto, vanno ad integrare esse stesse una componente "concorrenziale" rispetto al trattamento posto in essere dai singoli titolari del trattamento, ben potendo orientare (nell'ipotesi in cui il trattamento acceda ad un'offerta di beni o servizi) le scelte di quanti vedono le informazioni a sé riferite coinvolte nel trattamento (così dando attuazione al diritto all'autodeterminazione informativa), non diversamente dalle informazioni contenute sulle etichette e dai documenti informativi che i consumatori consultano prima di procedere all'acquisto di beni di consumo.

L'informazione rappresenta del resto il pre-requisito per un valido consenso al trattamento dei dati (che, per l'appunto, si vuole informato), ove lo stesso sia necessario. Consenso al trattamento, va qui ribadito, che non comporta alcuna "cessione" di dati personali, neanche quando acceda alla fruizione di servizi "gratuiti"; il diritto alla protezione dei dati personali, infatti, consiste anzitutto nel potere

[118] Cfr. Gruppo Art. 29, Parere 6/2014 sul concetto di interesse legittimo del responsabile del trattamento ai sensi dell'articolo 7 della direttiva 95/46/CE adottato il 9 aprile 2014 - WP 217.

dell'interessato di controllare l'uso che dei dati personali a sé riferiti viene fatto in relazione alle finalità per le quali i dati sono (legittimamente) trattati[119]. Anche in relazione ai *Big Data*, deve allora ribadirsi che la prospettiva della *commodification* dei dati personali non trova spazi nella cornice normativa eurounitaria[120]: e non solo muovendo dall'assunto (che pure da più parti si vorrebbe svalutare) della natura di diritto fondamentale del diritto alla protezione dei dati personali, ma perché puntuali indici normativi escludono la logica puramente appropriativa in relazione allo statuto giuridico dei dati personali. In questa prospettiva depongono infatti la libera revocabilità del consenso da parte dell'interessato al trattamento dei dati che lo riguardano, così come i diritti riconosciuti all'interessato, ivi compreso (come pure correttamente evidenziato nel corso delle audizioni) il nuovo diritto alla portabilità dei dati[121]: situazioni giuridiche che, complessivamente prese, vanno a comporre il diritto alla protezione dei dati personali.

E considerazioni non diverse possono trarsi da quanto espressamente contenuto nel considerando 24 della Direttiva (UE) 2019/770 relativa a determinati aspetti dei contratti di fornitura di contenuto digitale e di servizi digitali[122], come pure in ragione della dichiarata prevalenza dei principi espressi nel RGPD rispetto alla menzionata direttiva 2019/770 (cfr. artt. 3, par. 8; 16, par. 2)[123].

Questi indici normativi hanno trovato espressione anche nell'importante pronuncia della Corte di giustizia (Grande Sezione) la quale, in occasione del riconoscimento del cd. diritto all'oblio (sentenza del 13 maggio 2014 nella causa C-131/12, *Google Spain*), ha affermato che "*i diritti fondamentali* [nel caso di specie, proprio il diritto alla protezione dei dati personali] *prevalgono, in linea di principio, non soltanto sull'interesse economico del gestore del motore di ricerca, ma anche sull'interesse di tale pubblico a trovare l'informazione suddetta in occasione di una ricerca concernente il nome di questa persona*" (par. 97).

---

[119] Stefano Rodotà nell'esperienza italiana mise in risalto questa caratteristica coniando la formula (fortunata) "dal segreto al controllo", per evidenziare così il salto di qualità dalla vecchia *privacy* al nuovo diritto fondamentale alla protezione dei dati fondamentali. La Corte costituzionale tedesca fece propria, senza più abbandonarla dal 1983, la formula del "diritto all'autodeterminazione informativa".

[120] V. da ultimo anche le *Guidelines 2/2019 on the processing of personal data under Article 6(1)(b) GDPR in the context of the provision of online services to data subjects* Version 2.0 8 October 2019, in https://edpb.europa.eu/sites/edpb/files/files/file1/edpb_guidelines-art_6-1-b-adopted_after_public_consultation_en.pdf. In particolare al punto 54 si afferma che "*Considering that data protection is a fundamental right guaranteed by Article 8 of the Charter of Fundamental Rights, and taking into account that one of the main purposes of the GDPR is to provide data subjects with control over information relating to them, personal data cannot be considered as a tradeable commodity. Even if the data subject can agree to the processing of personal data,29 they cannot trade away their fundamental rights through this agreement*".

[121] Cfr. al riguardo il testo dell'Audizione del Prof. Mantelero (21 novembre 2017), p. 6, il quale aggiunge che, "*stante la collocazione della tutela dei dati personali nell'alveo dei diritti della personalità, la crescente commercializzazione di tali attributi non può snaturarne l'essenza: essi permangono strettamente inerenti la persona e non si assiste ad una reificazione degli stessi, che rimangono tutelati attraverso un modello basato su diritti individuali e sulla responsabilità di chi gestisce tali attributi (i.e. il titolare del trattamento nel caso dei dati personali). Da qui l'impossibilità di applicare logiche meramente proprietarie ai dati personali o di seguire modelli d'oltreoceano favorevoli all'assimilazione delle informazioni personali all'ambito della proprietà intellettuale*".

[122] "*La fornitura di contenuti digitali o di servizi digitali spesso prevede che, quando non paga un prezzo, il consumatore fornisca dati personali all'operatore economico. Tali modelli commerciali sono utilizzati in diverse forme in una parte considerevole del mercato. Oltre a riconoscere appieno che la protezione dei dati personali è un diritto fondamentale e che tali dati non possono dunque essere considerati una merce, la presente direttiva dovrebbe garantire che i consumatori abbiano diritto a rimedi contrattuali, nell'ambito di tali modelli commerciali*" […].

[123] Cfr. https://www.consilium.europa.eu/it/press/press-releases/2019/04/15/eu-adopts-new-rules-on-sales-contracts-forgoods-and-digital-content/.

Se da tempo il tema dei *Big Data* è stato posto all'attenzione dell'opinione pubblica nell'ambito dell'attività di comunicazione istituzionale del Garante, come pure nel corso di conferenze dedicate al tema[124], si registrano altresì alcune fattispecie nelle quali l'Autorità si è soffermata con propri provvedimenti sulla tematica qui considerata. Salvo tornare sul tema dell'anonimizzazione dei dati, può qui già menzionarsi il Parere reso sul PSN 2014-2016 (aggiornamento 2015-2016)[125], relativo all'elaborazione, in via sperimentale e a fini statistici, di informazioni di telefonia mobile con ulteriori microdati amministrativi e statistici provenienti dal Psn: ciò al fine di effettuare una stima dei flussi di mobilità intercomunali a livello aggregato e non individuale. A conferma quindi che un gioco a somma positiva tra protezione dei dati e *Big Data* è certo realizzabile, merita qui rilevare che in tale circostanza il Garante ha reputato idonee le assicurazioni fornite in ordine alla circostanza che presso il gestore telefonico i dati fossero raccolti in forma anonima[126].

4.8. *Big Data*, dati personali e procedure di anonimizzazione

Per quanto sia emerso nel corso dell'Indagine l'assunto secondo il quale le tecniche di *Big Data* sovente non richiedono l'uso di dati personali, è tuttavia necessario che chi intenda effettuare operazioni di trattamento secondo tale metodologia si accerti, in via preliminare, della natura personale o meno dei dati trattati, così da identificare (come si è visto) la cornice normativa di riferimento all'interno della quale opera. In questa prospettiva, sebbene la linea di demarcazione tra dati di natura personale e non possa essere in concreto difficile da tracciare – in particolare in ragione della possibilità di riconnettere informazioni apparentemente anonime (o anonimizzate) a individui singoli a seguito delle peculiari operazioni di trattamento effettuate (nel tempo sempre più agevolmente realizzabili, sia per le aumentate capacità di calcolo, sia per la pluralità di archivi in ipotesi utilizzabili, aventi anche genesi ed utilizzi prospettici diversi al tempo della raccolta) –, un utile contributo può essere tratto dalle decisioni delle autorità di protezione dei dati e dagli indirizzi assunti dal Comitato europeo per la protezione dei dati, come pure dalle migliori prassi via via elaborate (e comunque soggette a continui aggiornamenti) in tema di anonimizzazione dei dati personali.

Comprensibilmente (ed opportunamente) una via d'uscita (per molti quella maestra) viene ricercata nelle soluzioni tecnologiche, prima fra tutte nella effettiva anonimizzazione dei dati che dovrebbero andare a costituire i *Big Data*[127], cui associare, ove necessario, misure di natura organizzativa e/o contrattuale preordinate allo stesso scopo.

---

[124] Tra queste, si pensi ai contributi raccolti nel volume curato dal Garante, *Big Data e Privacy*, Atti del Convegno del 30 gennaio 2017, Roma, 2017.

[125] Provv. 18 settembre 2014, n. 411 (doc. web n. 3458502), con il quale si è valutato il lavoro statistico denominato "*Uso a fini statistici dei Big Data*" (identificato con il codice IST-02589).

[126] V. pure, in tempi più recenti, a conferma dell'incremento delle fattispecie di impiego di tecniche *big data*, il Parere sullo schema di Programma statistico nazionale 2017-2019, Aggiornamento 2018-2019, provv. 9 maggio 2018, n. 271, doc. web n. 9001732).

[127] In questa prospettiva si è affermato che "*l'accesso ai dati deve essere regolamentato, secondo necessità e proporzionalità: libera fruizione degli open data (rendere i dati visualizzabili e utilizzabili per tutti) ma anche rispetto della privacy (attraverso norme e policy sull'utilizzo di questi dati). Ad esempio, un tema importante è quello dell'anonimizzazione dei dati: i dati anonimizzati sono meno pericolosi dal punto di vista della privacy rispetto a quelli che non lo sono, e rimangono un problema tecnologico, ovvero di competenze. Per anonimizzare un dato è necessario un algoritmo, cioè una serie di linee di codice di software che deve essere continuamente aggiornato, che non permetta il reverse engineering, la reingegnerizzazione del processo, garantendo un'architettura a sicurezza elevata. Quindi, anonimizzazione, necessità e proporzionalità dell'utilizzo dei dati sono, di fatto, problemi che si risolvono tecnologicamente, certo con una policy opportuna alle spalle ma di forte impronta tecnologica perché il processo possa*

E tuttavia, come si è accennato, anche questa strada deve essere percorsa con prudenza, con analisi casistica, avendo da tempo la comunità scientifica, come pure le autorità di protezione dei dati[128] evidenziato i rischi di re-identificazione degli interessati utilizzando *dataset* ulteriori (pur privi di identificativi individuali)[129]; rischio amplificato dalla (via via) crescente massa di informazioni liberamente disponibili (anche per il legittimo riuso) *on-line*.

Pari attenzione deve essere prestata nel caso in cui l'utilizzo di *Big Data* si incentri sul trattamento di *dataset* acquisiti presso terzi: il titolare del trattamento, oltre a considerare l'incrementato rischio di re-identificazione che potrebbe derivarne, deve altresì accertarsi della sussistenza delle condizioni di riutilizzo dei dati così acquisiti (circostanza peraltro ricorrente in ambiti più tradizionale, quali quelli legati alla circolazione di *database* contenenti dati personali per finalità di *marketing*)[130].

Al netto di tali considerazioni, le tecniche di analisi basate sul paradigma *Big Data* comportano una serie di rischi diretti o indiretti che è necessario fronteggiare con misure di sicurezza adeguate, efficaci, realizzate a regola d'arte (*state of the art*) e continuamente valutate e aggiornate, sia nell'ottica della conformità all'art. 25 del RGPD (*data protection by design/default*) che del rispetto dell'art. 32 del RGPD (sicurezza dei trattamenti). È infatti evidente come le elaborazioni *Big Data*, pur basate su *dataset* anonimi o (semplicemente) ritenuti anonimi, rechino pericoli di possibile pregiudizio a diritti e libertà degli interessati cui i dati possono essere riferiti: come accennato, pur partendo da dati anonimi (perché ottenuti da dati personali sottoposti a procedure di mascheramento o di anonimizzazione) molto spesso permane nei dati di *output* la possibilità che si producano effetti di *singleouting* o di reidentificazione, con conseguenze che si riflettono su un individuo o un gruppo di individui, ancorché non compiutamente individuati.

Esempi di *singleouting* ricorrono infatti nella letteratura statistica e informatica, e una casistica ricorrente riguarda le elaborazioni possibili su dati demografici e sanitari, possibilmente correlati ad altre raccolte pubblicamente disponibili, come le liste elettorali. Alla fine degli anni '90 del secolo scorso la ricercatrice americana Latanya Sweeneyche contribuì, insieme alla collega Pierangela Samarati[131], a gettare le basi della teoria della *k-anonymity* quale misura per limitare il potenziale

avvenire in maniera certa e completa": cfr. Piacentini, Intervento al Convegno *Big Data e Privacy. La nuova geografia dei poteri*, cit., p. 70.

[128] Cfr. Gruppo "articolo 29", parere 4/2007 sul concetto di dati personali, adottato il 20 giugno 2007 (WP 136); Parere 6/2013 sui dati aperti e sul riutilizzo delle informazioni del settore pubblico ("ISP"), adottato il 5 giugno 2013 (WP207); Parere 05/2014 sulle tecniche di anonimizzazione adottato il 10 aprile 2014 (WP 216), che si sofferma, in particolare, su tecniche quali l'aggiunta del rumore statistico, le permutazioni, la privacy differenziale, l'aggregazione, il k-anonimato, la l-diversità e la t-vicinanza.

[129] V. in merito la trattazione di G. D'Acquisto e M. Naldi, Big Data *e* Privacy by design, Torino, 2017, *passim*; ENISA (a cura di), *Privacy by design in big data. An overview of privacy enhancing technologies in the era of big data analytics*, realizzato da G. D'Acquisto – J. Domingo-Ferrer – P. Kikiras – V. Torra – Y.-A. de Montjoye – A. Bourka, 2015, in https://www.enisa.europa.eu/activities/identity-and-trust/library/deliverables/big-dataprotection; la circostanza è evidenziata da ultimo anche da E. Palmerini, Dalle *smart cities* allo *scoring* del cittadino, in Atti del Convegno *I Confini del Digitale. Nuovi scenari per la protezione dei dati*, cit., p. 20, 28.

[130] Cfr., ad es., con riguardo alla sussistenza del requisito del consenso degli interessati, Garante, Linee guida in materia di attività promozionale e contrasto allo spam, provv. 4 luglio 2013, doc. web n. 2542348, punto 2.6.3.

[131] Samarati, Pierangela; Sweeney, Latanya (1998). "*Protecting privacy when disclosing information: k-anonymity and its enforcement through generalization and suppression*". Harvard Data Privacy Lab.

identificativo delle grandi raccolte informatizzate di dati, e pubblicò i primi articoli scientifici che suscitarono interesse e preoccupazione nel pubblico[132].

Il nuovo Regolamento europeo introduce, anche rispetto agli aspetti di sicurezza dei trattamenti, diverse novità e tra queste merita di soffermarsi sul richiamato art. 32 sulla sicurezza dei trattamenti, in cui viene posto l'accento sull'esigenza della protezione dei dati come via per garantire i diritti e le libertà delle persone, e viene suggerita, tra le altre, la misura della *pseudonimizzazione*, procedura meno incisiva della *anonimizzazione* ma idonea ad abilitare alcune forme di trattamento riducendo i rischi di reidentificazione.

C'è da osservare, però, che il confine netto, la dicotomia tra dato anonimo, e perciò non personale, e dato personale, non corrisponde alla realtà concreta dei trattamenti. Si riscontra piuttosto una sorta di continuità tra questi due concetti, che comporta la gradazione progressiva dall'anonimato impersonale all'identificazione: una dose residua di identificabilità è quindi presente in tante raccolte di dati comunemente ritenute anonime, mentre è individuabile anche quantitativamente (con metodi matematici) un *trade-off* tra anonimizzazione e utilità del dato. La tecnologia può rendere oggi un dato anonimizzato e domani renderlo nuovamente dato personale. D'altra parte, la definizione stessa di dato personale, valorizzando la potenzialità del *linkage* tra differenti dati, anche in possesso e sotto il controllo di differenti soggetti, nella qualificazione di un dato come dato personale, fa sì che la nozione di dato personale acquisti un'ampiezza che sfugge a molti.

Un dato di fatto incontrovertibile è che la disponibilità crescente di dati rende le persone sempre più identificabili. Questo rischio non può essere affrontato con strumenti analitico-matematici ma può essere più efficacemente affrontato con *policy* che comprendano considerazioni etiche e valutazioni su chi siano i destinatari dei dati anonimizzati, quali siano le garanzie di correttezza che essi offrono, che possibilità di *disclosure* successiva permane nei dati conferiti.

Venendo allora a considerare più da vicino il tema dell'anonimizzazione dei dati, va ricordato che obiettivo di tale operazione di trattamento è impedire che sia possibile, utilizzando mezzi "ragionevoli": 1) isolare una persona in un gruppo; 2) collegare un dato anonimizzato a dati riferibili a una determinata persona censiti in una differente base di dati; 3) la deduzione di nuove informazioni personali a partire da un dato anonimizzato.

Diverse sono le tecniche di anonimizzazione che nel contesto dei *Big Data* possono essere adottate in base alle esigenze di trattamento, e ricadono nelle macrocategorie della randomizzazione e della generalizzazione. Le tecniche di randomizzazione e quelle di generalizzazione mirano a far diminuire le probabilità di successo dei tentativi di ricondurre a una persona individuata i dati sottoposti ad anonimizzazione e forniscono delle metriche per rendere quantificabile e matematicamente limitabile il rischio residuo.

La randomizzazione può consistere nell'introduzione di un "rumore" statistico o nello *shuffling* degli attributi riferibili a un insieme di individui (sostanzialmente permutando le proprietà di un membro di un insieme di individui con quelle di altri appartenenti allo stesso gruppo).

Le tecniche di generalizzazione sono invece basate sull'idea, presente in materia censuale e demografica, di evitare effetti di *singleouting* modificando la scala di rappresentazione di determinati

[132] L. Sweeney, *Simple Demographics Often Identify People Uniquely*. Carnegie Mellon University, Data Privacy Working Paper 3. Pittsburgh 2000. [https://dataprivacylab.org/projects/identifiability/paper1.pdf]

attributi: per esempio, in un *dataset* demografico si può scegliere di utilizzare, al posto del CAP (codice di avviamento postale che, unito ad altre informazioni anagrafiche permette con elevata probabilità di arrivare al *single out*), un codice geografico su più ampia scala che trasformi il riferimento a una ristretta area geografica (quartiere, città) in un dato riferibile a una molteplicità di luoghi (provincia, regione…), con ciò abbattendo drasticamente la probabilità di reidentificare i soggetti cui si riferiscano i dati elaborati.

Procedure analoghe possono essere messe in atto per i riferimenti temporali (minore dettaglio sulle date, con riferimento, per esempio, all'anno di nascita o a intervalli di date, e non alla data esatta).

Con queste tecniche si realizza per via informatica una protezione "per affollamento" (*crowding*), analoga a quella dell'individuo che cerca di mischiarsi a una folla indistinta di persone per rendersi meno riconoscibile in pubblico.

Le tecniche di anonimizzazione devono però produrre dati non privi di valore a fini di elaborazione, altrimenti verrebbe meno l'utilità di questo procedimento o di altri analoghi basati sullo stesso paradigma. Per esempio, se scopo di un'elaborazione su larga scala fosse uno studio epidemiologico, l'ampliamento dell'area geografica attribuibile a un record individuale rischierebbe, se troppo marcato, di annullare il valore informativo relativo alla maggiore o minore pertinenza di una certa patologia a un determinato territorio o area.

La valutazione sul rischio potenziale connesso alla disponibilità di dati anonimizzati dipende però da informazioni che non necessariamente potranno essere disponibili in maniera sufficiente al momento della decisione e della scelta del metodo di anonimizzazione: una sorgente pubblica di dati potrebbe essere resa disponibile successivamente alla pubblicazione di un *dataset* ritenuto anonimo che, in modo imprevedibile, si vedrebbe ricongiungibile a informazioni, non note *a priori,* che consentirebbero il *linkage* e il *singleout* ai danni di determinate persone. L'anonimizzazione non può pertanto essere considerata un'operazione una tantum e i relativi rischi dovrebbero essere oggetto di un riesame periodico da parte dei titolari del trattamento[133].

La pseudonimizzazione, invece, è una misura meno radicale dell'anonimizzazione, prevista dal nuovo regolamento europeo quale misura di sicurezza utilizzabile, su valutazione del titolare, per determinati tipi di trattamento. In particolare, con la pseudonimizzazione viene mantenuta la corrispondenza 1:1 del dato pseudonimizzato con il dato originario. I *dataset* pseudonimizzati recano intatto il valore informativo statistico, non essendo frutto di alterazioni di scala o di distorsioni di varia natura (come nei metodi di randomizzazione). La pseudonimizzazione quindi, nelle varie forme in cui può essere realizzata, consente di mantenere l'utilizzabilità statistica dei dati, senza annullarne il valore informativo, tutelando nel contempo l'identità dei soggetti cui si riferiscono.

Sarà responsabilità del titolare che ricorre alla pseudonimizzazione quale misura di sicurezza (indicata, a titolo esemplificativo, dall'articolo 32 del RGPD) curare il mantenimento di quella cesura tra il dato in chiaro e il dato pseudonimizzato, per evitare il più possibile la connessione dei dati alle persone cui si riferiscono.

Esempi di tecniche di pseudonimizzazione comprendono il ricorso a procedure di *hashing*, ovvero a funzioni matematiche non invertibili che associano a un insieme di dati una stringa di caratteri che non ha relazione alcuna con il contenuto e la semantica del dato, al di là della corrispondenza

[133]Cfr. Gruppo "articolo 29", Parere 05/2014 sulle tecniche di anonimizzazione adottato il 10 aprile 2014 (WP 216), p.4.

matematica, e che possono essere rafforzate dal ricorso congiunto a *chiavi di hashing* che rendano più netta la cesura tra il dato originario e il dato "*hashed*" anche ricorrendo ad algoritmi di *hashing* di pubblico dominio, come MD5 o SHA-256, limitando le possibilità di applicare metodi di reidentificazione per *matching*.

Il dato pseudonimizzato rimane pur tuttavia *dato personale* e, come tale, soggetto agli obblighi di protezione sanciti dal regolamento. Qual è dunque il vantaggio del ricorso alla pseudonimizzazione? In primo luogo essa può costituire una salvaguardia nel caso in cui le misure tecniche di sicurezza non siano riuscite a proteggere i dati da una violazione: l'incentivo per i titolari a usare la pseudonimizzazione (in luogo di dati direttamente identificabili) deriva essenzialmente dal differente regime sanzionatorio, ad esempio nel caso di incidenti di sicurezza.

La tecnica più promettente per ridurre il rischio di re-identificazioni è oggi la *Differential Privacy*, che offre il maggior numero di tutele per gli interessati integrando i benefici delle tecniche di generalizzazione e randomizzazione, in quanto prevede un meccanismo di accesso ai dati basato su interrogazioni (*query based mechanism*) e non sulla pubblicazione di dati aggregati o randomizzati (*sanitized data*), ed è molto robusta rispetto alla possibilità di impiegare informazioni ausiliarie (anche pubblicamente disponibili) per la re-identificazione[134].

Conclusivamente, deve ritenersi che chi intenda avvalersi dei *Big Data* facendo uso di tecniche di anonimizzazione è tenuto comunque ad effettuare periodicamente un *assessment* approfondito circa il rischio di re-identificazione, al fine di valutare la "robustezza" delle metodologie impiegate per procedere all'anonimizzazione dei dati e documentando il processo seguito[135].

4.9. *Big Data* e principio di finalità

Né, sotto un diverso profilo, l'anonimizzazione dei dati può rappresentare un *escamotage* per effettuare trattamenti non compatibili con le finalità originarie della raccolta: in questo senso si è di recente espresso il Garante con riguardo al tema del possibile trattamento dei dati contenuti nelle fatture elettroniche da parte di operatori privati, operanti in qualità di responsabili del trattamento, che volessero usarli per autonome (e non meglio precisate) elaborazioni "statistiche" [136].

---

[134] Si segnala, ad esempio, l'adozione di un meccanismo di *Differential Privacy* da parte di Apple nella versione iOS 10 del proprio sistema operativo http://appleinsider.com/articles/16/06/20/apples-differential-privacy-analyzes-the-group-protects-the-individual.

[135] Considerazioni di analoga natura sono state di recente svolte dal Garante nel Parere sullo schema di Programma statistico nazionale 2017-2019, Aggiornamento 2018-2019 (provv. 9 maggio 2018, n. 271, doc. web n. 9001732) con riferimento all'introduzione, in molti lavori statistici (ed in particolare quelli afferenti al Sistema dei registri), di una nuova variabile denominata "codice univoco indirizzo (CUI) di residenza SIM" o "codice univoco indirizzo (CUI) di domicilio SIM". Al riguardo, il Garante ha segnalato che l'utilizzo di tali codici, riferibili alle persone fisiche che risultano, di volta in volta, attraverso la mera consultazione di pubblici registri, intestatarie, proprietarie, residenti o titolari di attività economiche nei luoghi individuati, è idoneo a rivelare, con tecniche di *linkage* e di georeferenziazione degli indirizzi, sia i luoghi di dimora abituale, di lavoro, di studio, di abitazione e di cura, sia i legami tra individui, luoghi, enti e istituzioni, aumentando, così, esponenzialmente il patrimonio informativo riferibile all'intera popolazione e, quindi, anche i connessi rischi per le libertà e i diritti degli interessati. Anche in relazione ad altre sperimentazioni che prevedevano l'utilizzo di fonti di telefonia mobile, nel medesimo parere si è ribadito che l'utilizzo di queste informazioni comporta specifici rischi per la riservatezza e la protezione dei dati personali degli interessati, tenuto anche conto che, grazie alle nuove tecnologie e alle nuove tecniche di analisi, elaborazione e interconnessione dei dati, risulta spesso possibile (o, comunque altamente probabile) la re-identificazione di un interessato anche attraverso informazioni apparentemente anonime (c.d. "*single-out*")..

[136] Cfr. provv. 20 dicembre 2018, n. 511, doc. web n. 9069072: con tale provvedimento, esaminati alcuni modelli contrattuali utilizzati dalle maggiori società produttrici di software gestionale e fiscale – in cui sono presenti alcune

Il principio di finalità rappresenta uno dei principali baluardi per il rispetto del diritto alla protezione dei dati personali e per salvaguardare la dignità dell'individuo. Ancorché esso non costituisca una barriera insormontabile rispetto ai trattamenti con le tecniche *Big Data*, impone tuttavia che chi intenda avvalersene debba procedere ad una approfondita valutazione di compatibilità con le finalità che originariamente hanno determinata la raccolta dei dati [137].

Interrogativi devono inoltre essere sollevati circa la liceità del "farsi" dei *Big Data*, anche in ordine alla possibilità di trattare, sulla scorta di una legittima base giuridica, i *dataset* oggetto di analisi, come pure del loro utilizzo, atteso che ciò comporta per lo più in un uso "secondario" dei dati che deve quindi misurarsi con il principio di compatibilità del trattamento in relazione alla finalità rispetto alle ragioni che hanno determinato l'originaria raccolta (art. 5, par. 1, RGPD).

Anche con riguardo ai dati presenti in internet o tratti dai *social media*, la mera circostanza che gli stessi siano resi pubblici (e quindi agevolmente "processabili") non giustifica un loro riutilizzo con tecniche *Big Data*, in particolare rispetto a finalità che nulla hanno a che vedere con quelle per le quali tali informazioni erano state conferite[138].

4.10. *Big Data*, principi di qualità e minimizzazione dei dati

Del pari determinante è il rispetto del principio di qualità – che pure deve tradursi in un approccio *by design* in ogni stadio del processo, vale a dire all'introduzione di procedure volte ad assicurare sistematicamente e su larga scala la qualità dei dati raccolti e l'appropriatezza dei criteri di analisi dei

---

clausole suscettibili di violare, in particolare, gli artt. 5, 6 e 28 del Regolamento, che evidenziano elevati rischi di utilizzi impropri dei dati personali nell'ambito dei trattamenti effettuati dagli intermediari e dagli altri soggetti delegati dagli operatori economici nel processo di fatturazione, non solo con riferimento a trattamenti illeciti, ma anche alla proliferazione di possibili collegamenti e raffronti tra fatture di migliaia di operatori economici – non si è ritenuta conforme al Regolamento la clausola secondo cui una società produttrice di software gestionale e fiscale possa procedere "*all'elaborazione e utilizzo di informazioni puramente statistiche, su base aggregata e previa anonimizzazione, raccolte in relazione all'utilizzo dei Servizi da parte del Cliente e del Terzo Beneficiario, ivi incluse informazioni relative ai meta-dati associati ai Documenti, a fini di studio e statistici, concedendo a tal fine [alla predetta società] una licenza non esclusiva, perpetua, irrevocabile, valida in tutto il mondo e a titolo gratuito, ad utilizzare tali informazioni per dette finalità*". Si è al riguardo ribadito che i dati personali contenuti nelle fatture non sono riferiti esclusivamente all'operatore economico che le ha emesse e ricevute, ma pure ai terzi – anche persone fisiche – con cui intrattiene rapporti economici. I trattamenti svolti in qualità di responsabile e di subresponsabile devono pertanto essere limitati solo ed esclusivamente a quanto necessario per la fornitura dei servizi forniti ai titolari e, dunque, per l'esecuzione del contratto stesso, senza introdurre operazioni di trattamento ulteriori (ivi compresa l'anonimizzazione dei dati) preordinate al perseguimento di finalità proprie del responsabile, rispetto alle quali deve essere, di volta in volta, valutata la rispondenza ai requisiti del Regolamento, quali, in particolare, i presupposti di liceità del trattamento e il rispetto dei principi applicabili al trattamento dei dati personali.

[137] V. al riguardo Article 29 Data Protection Working Party, WP 203, *Opinion 03/2013 on purpose limitation*, Adopted on 2 April 2013, https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2013/wp203_en.pdf, p. 35 e, con indicazioni di maggior dettaglio, 45 ss..

[138] V. provv. 18 aprile 2019, n. 96, doc. web n. 9105201, punto 5, E), in materia di propaganda elettorale e comunicazione politica; cfr. già, in merito, provv. Garante, 11 gennaio 2001, doc. web n. 40823, nel quale si è stabilito che "*la conoscenza di fatto degli indirizzi che si realizza in tali casi non può essere disgiunta dalla finalità per cui essa avviene. Contrasta, pertanto, con i principi di correttezza e finalità del trattamento raccogliere i dati che singoli utenti "lasciano" in un newsgroup, forum, ecc. solo per le finalità di specifica discussione su determinati temi, hobbies, ecc., ed utilizzarli per altri scopi che nulla hanno a che vedere – anche indirettamente – con l'argomento per il quale l'utente partecipa ad una discussione più o meno "pubblica" ed indica il proprio recapito e le proprie generalità*". Ed ancora è stato ribadito con il provv. 30 dicembre 2002, doc. web n. 1067376, che "*i dati personali disponibili in rete in relazione ad eventi e delimitate finalità non sono liberamente utilizzabili per l'invio generalizzato di e-mail aventi contenuto commerciale o pubblicitario*". Con riguardo alla vicenda Cambridge Analytica, v. European Data Protection Board, *Statement 2/2019 on the use of personal data in the course of political campaigns*, Adopted on 13 March 2019, https://edpb.europa.eu/sites/edpb/files/files/file1/edpb-2019-03-13-statement-on-elections_en.pdf.

dati selezionati – con riguardo ai dati che possono formare oggetto di trattamento nei *Big Data*. Il trattamento di informazioni di scarsa qualità (ad es. incomplete, inesatte o risalenti) come pure l'impiego di tecniche di ricerca (algoritmi) che presentano errori di configurazione, sono suscettibili di condurre ad inferenze scorrette che si ripercuotono, in ultima analisi, sul processo (conoscitivo o decisionale) posto in essere e quindi su singoli o gruppi (più o meno ampi) di individui.

Prima dell'impiego su larga scala dei risultati generati tramite i *Big Data* è preferibile effettuare sperimentazioni e verifiche della correttezza degli esiti su scala ridotta. La individuazione di vizi (bias) sistemici nei dati oggetto di trattamento con i *Big Data* o nei processi valutativi (o predittivi) realizzati loro tramite deve condurre all'immediata revisione dei modelli utilizzati e all'adozione delle necessarie misure correttive. Il mancato approntamento di tali misure, preventive o successive, rispetto all'utilizzo dei *Big Data*, rileva dal punto di vista della complessiva correttezza del trattamento.

Tra gli aspetti che possono presentare profili di particolare criticità va annoverato il principio di minimizzazione dei dati, che può riverberare sia dal punto di vista quantitativo e qualitativo dei dati utilizzati (ad es. escludendo, ogniqualvolta possibile, l'uso di dati sensibili), sia dal punto di vista della profondità temporale dei dati utilizzati (valutando adeguatamente, quindi, i tempi di conservazione degli stessi). Con riguardo ad entrambi questi profili, ancora una volta, si tratta di considerare in concreto gli ambiti nei quali i trattamenti con tecniche *Big Data* possono essere effettuate, potendosi giustificare trattamenti di dati maggiormente "voluminosi" e risalenti nel tempo in taluni ambiti (ad esempio, nel contesto della ricerca medico scientifica) rispetto ad altri (ad esempio, in ambito commerciale). Queste varie circostanze devono formare oggetto di valutazione, anche per il tramite della menzionata valutazione d'impatto, e contribuiranno a definire il *design* del trattamento al fine di minimizzare i rischi di violazione dei diritti degli interessati.

Nel definire il corretto *design* che consente il trattamento di informazioni mediante *Big Data* può altresì essere valorizzata la previsione contenuta nell'art. 11, par. 1, del RGPD, secondo la quale "*se le finalità per cui un titolare del trattamento tratta i dati personali non richiedono o non richiedono più l'identificazione dell'interessato, il titolare del trattamento non è obbligato a conservare, acquisire o trattare ulteriori informazioni per identificare l'interessato al solo fine di rispettare il presente regolamento*".

4.11. *Big Data*, valutazione d'impatto *privacy* e *accountability*

Al fine di minimizzare i rischi di violazione dei diritti individuali per il tramite dei *Big Data*, considerate le caratteristiche dei trattamenti in esame (che possono determinare anche l'adozione di decisioni individuali sulla base di analisi "predittive"), è quindi necessario adottare misure preventive e processi interni volti a commisurare il rischio sui diritti degli interessati (che sarà tanto più elevato quanto più intrusivi nella sua sfera personale potranno essere gli effetti dell'analisi effettuata). In questa prospettiva si colloca, peraltro, la valutazione d'impatto sulla protezione dei dati prevista dall'art. 35 del RGPD, cui, con alta probabilità, devono essere sottoposti i trattamenti di dati posti in essere con la tecnica dei *Big Data*, in particolare con riguardo ai casi in cui il trattamento comporti "*una valutazione sistematica e globale di aspetti personali relativi a persone fisiche, basata su un trattamento automatizzato, compresa la profilazione, e sulla quale si fondano decisioni che hanno effetti giuridici o incidono in modo analogo significativamente su dette persone fisiche*" (cfr. art. 35, par. 3, lett. a), del RGPD). Essa può consentire non solo di individuare possibili rischi per i diritti dei singoli ma anche di evidenziare possibili conseguenze indesiderate di natura etica o sociale (anche

per gruppi di individui); inoltre, se correttamente effettuata – ove designato, sentito il responsabile per la protezione dei dati –, può condurre all'adozione di misure tali da eliminare o comprimere i rischi di violazione dei diritti di individui e gruppi ovvero ad individuare gli ambiti da sottoporre alla consultazione preventiva dell'autorità di controllo (cfr. art. 36 del RGPD).

Come evidenziato nelle richiamate *Guidelines on the protection of individuals with regard to the processing of personal data in a world of big data* del Consiglio d'Europa può risultare altresì utile (ed è certamente una misura che può rientrare a pieno titolo nel concetto di *accountability*) coinvolgere nel corso della valutazione i soggetti o le categorie di soggetti che potrebbero essere interessati dai trattamenti effettuati.

Nella prospettiva dell'*accountability* e della trasparenza nei confronti degli interessati, salvaguardate informazioni di natura riservata, la valutazione d'impatto potrebbe altresì essere resa pubblica dal titolare del trattamento.

Una piena partecipazione degli interessati al trattamento effettuato con tecniche di *Big Data*, vale a dire in presenza di un consenso pienamente informato e liberamente manifestato (si pensi ad esempio a trattamenti nel contesto sanitario, anche con la partecipazione di rappresentanti degli interessati), costituisce la soluzione preferibile per assicurare un pieno rispetto dei principi di protezione dei dati personali e soddisfare le esigenze di *accountability* introdotte dal RGPD. Data la natura dei trattamenti in parola, tale processo può risultare difficoltoso, ma approcci innovativi volti a favorire processi di partecipazione individuale, anche mediante un più intenso uso delle ICTs come canale di comunicazione con i soggetti interessati (o con gli esponenti dei portatori degli interessi collettivi coinvolti da tali operazioni di trattamento: consumatori, risparmiatori, pazienti, ecc.) dovrebbero essere esplorati (come accaduto, ad esempio, nel settore della ricerca genetica e delle biobanche). Diversamente, in assenza di puntuali previsioni legislative, quanto più un trattamento con tecniche di *Big Data* viene effettuato in assenza della partecipazione dell'interessato (se non della sua conoscenza), tanto più aumentano i rischi di violazione delle discipline di protezione dei dati personali.

L'*accountability* del titolare del trattamento non si esaurisce nella fase "progettuale" nell'utilizzo dei *Big Data*; opportune misure devono essere poste in essere per monitorare con sistematicità l'efficacia delle soluzioni predisposte, affinandole se necessario, e verificare la qualità del processo posto in essere.

4.12. *Big Data* e processi decisionali automatizzati

Non meno determinante è la fase terminale del processo intrapreso con tecniche basate sui *Big Data* al fine di prevenire, come richiesto dall'art. 22 del RGPD, l'adozione di decisioni automatizzate in capo a singoli individui, in grado di incidere significativamente sugli stessi (si pensi, ma sono soltanto esemplificazioni, a processi automatizzati in relazione alla valutazione del merito creditizio o alle politiche di selezione del personale).

In linea di principio, l'uso dei *Big Data* non dovrebbe comportare l'esclusione dell'intervento umano nel processo decisionale, il quale deve potersi dissociare dalla soluzione proposta all'esito del processo automatizzato. Tali aspetti devono essere presi in considerazione nella valutazione d'impatto.

Devono essere inoltre adottate dal titolare del trattamento le misure necessarie volte a fornire all'interessato coinvolto in un processo decisionale automatizzato informazioni significative sulla logica utilizzata (cfr. art. 13, par. 2, lett. f), del RGPD)[139].

4.13. *Big Data* e grandi archivi pubblici

Ancorché l'Indagine congiunta si sia prevalentemente incentrata sull'utilizzo dei *Big Data* nella dimensione delle reti di comunicazione elettronica e nel settore privato, tali ambiti non possono reputarsi esclusivi; le potenzialità di tale fonte informativa costituiscono, infatti, terreno propizio per chi reputa che l'utilizzo dei *Big Data* possa incidere anche sull'efficienza della pubblica amministrazione e, quindi, sui trattamenti di dati personali effettuati in ambito pubblico.

L'indeterminatezza delle finalità perseguite in concreto con i *Big Data*, i rischi di reidentificazione degli interessati e l'opacità delle logiche applicate al trattamento – aspetti tutti già ampiamente evidenziati nelle pagine che precedono – entrano in conflitto con i requisiti richiesti dalla normativa internazionale ed europea in materia di protezione dei dati personali a tutela dei diritti e libertà degli individui.

Le criticità del fenomeno sono ancora più evidenti nel settore pubblico, dove nemmeno il consenso può riequilibrare il rapporto tra l'interessato/cittadino e il potere pubblico.

Peraltro, non bisogna confondere l'utilizzo dell'immenso patrimonio informativo quale quello contenuto negli archivi progressivamente acquisiti e digitalizzati nel tempo dalle pubbliche amministrazioni per il perseguimento di finalità legittime e determinate, con le potenzialità insite nel fenomeno dei *Big Data*.

Già con riferimento all'integrazione, presso l'Istat, a fini statistici, di intere basi dati amministrative relative alla totalità della popolazione, con trattamenti automatizzati volti anche a definire il profilo o la personalità dell'interessato (non finalizzati a una ricaduta amministrativa sugli individui, vietata dalla predetta normativa sulla protezione dei dati personali), il Garante ha richiesto l'introduzione di uno specifico quadro di garanzie a tutela degli interessati, specie in relazione alla natura, alla qualità dei dati, alle modalità del trattamento, nonché alle misure di sicurezza.

Profili di criticità sono emersi anche con riferimento alle iniziative di più recente progettazione nell'ambito delle quali va evidenziata la riforma del CAD che recepisce all'art. 50-*ter* il *Data & Analytics Framework* (DAF), introducendo la Piattaforma Digitale Nazionale Dati (PDND). La creazione della suddetta Piattaforma comporta un accentramento e una duplicazione di tutti i dati detenuti dalle pubbliche amministrazioni per finalità del tutto generiche, realizzando di fatto una concentrazione presso un unico soggetto di informazioni, anche sensibili e sensibilissime, con evidenti rischi di vulnerabilità dei dati stessi ovvero di possibili usi distorti.

[139] In merito all'utilizzo di processi decisionali automatizzati da parte della pubblica amministrazione, facendo leva sui principi fondamentali dell'attività amministrativa di imparzialità, pubblicità e trasparenza, il Consiglio di Stato ha sottolineato l'esigenza (rispondente all'irrinunciabile necessità di poter sindacare come il potere sia stato concretamente esercitato, ponendosi in ultima analisi come declinazione diretta del diritto di difesa) che l'algoritmo a cui viene "delegata" la decisione amministrativa sia pienamente conoscibile e soggetto al pieno sindacato del giudice amministrativo. Il che comporta, tra le altre conseguenze, che la "formula tecnica", che di fatto rappresenta l'algoritmo, deve essere corredata da spiegazioni che la traducano nella "regola giuridica" ad essa sottesa e che la rendano leggibile e comprensibile, sia per gli interessati che per il giudice. In mancanza della totale trasparenza e conoscibilità dell'algoritmo, la procedura informatizzata è illegittima e il provvedimento finale da annullare (cfr. Consiglio di Stato, sez. VI, 8 aprile 2019, n. 2270).

La necessaria valorizzazione del patrimonio informativo pubblico non deve, infatti, avvenire a discapito della tutela dei diritti fondamentali e con possibili ricadute anche in termini di sicurezza nazionale.

Un trattamento fondato sui *Big Data* in ambito pubblico richiede un'idonea base legale del trattamento che assicuri ai cittadini, oltre alla trasparenza delle decisioni, la proporzionalità del ricorso *ex lege* a tale metodologia rispetto all'obbiettivo di interesse pubblico perseguito e l'individuazione, nel rispetto del principio di *privacy by design*, di adeguate garanzie da integrare nel trattamento, dopo aver accuratamente valutato i rischi elevati per i diritti e le libertà degli interessati.

Del pari interessati sono gli ambiti della sanità[140] (ed in prospettiva quello del cd. *mobile health*) e della ricerca, caratterizzantisi già per l'intensa digitalizzazione delle informazioni e che quindi, con maggiore probabilità, possono formare oggetto di "cattura" da parte dei soggetti in grado di svolgere le analisi proprie dei *Big Data* per individuare correlazioni (per lo più in via probabilistica) tra le malattie e le possibili (con)cause.

Da segnalare l'impiego sempre più diffuso della c.d. "tecnologia impiegata" alle grandi banche dati sanitarie detenute, per lo più, per fini di governo, da regioni e da enti centrali. L'"intelligenza aumentata" è vista come strumento per incrementare i risultati delle analisi tradizionali sui dati sanitari, al fine di individuare nuovi indicatori utilizzabili soprattutto nella politica sanitaria. Il perfezionamento di tali tecnologie richiede il trattamento di informazioni sanitarie reali, utili soprattutto per rappresentare l'evoluzione delle cronicità, che rappresentano una voce sempre più considerevole nel bilancio della spesa sanitaria. L'ordinamento giuridico si trova, pertanto, di fronte ad una nuova dimensione del trattamento dei dati, la cui estensione non si misura solo con le delicate problematiche tipiche dell'utilizzo dei *Big Data*, ma, come si è anticipato, anche con nuovi quesiti etico-giuridici[141].

Anche nell'ambito dei trattamenti effettuati per finalità di polizia le tecniche di *Big Data* (ivi comprese le ipotesi di *data scraping* in internet) sono sempre più frequentemente prese in considerazione, non senza che siano rivolti avvertimenti al rischio di conseguenze discriminatorie derivanti (in particolare) dai bias che possono caratterizzare l'algoritmo chiamato ad operare sui *dataset* interessati[142].

4.14. Prospettive

Le profonde implicazioni dei *Big Data* sulla società nel suo complesso (al di là degli effetti più o meno immediati che essi possono produrre sui singoli) che l'Indagine congiunta ha lasciato trasparire – e per ciò solo essa ha rappresentato un valore aggiunto – fanno sorgere più dubbi di quante risposte si possano, allo stato, dare.

---

[140] Cfr. European Commission, *Study on Big Data in Public Health, Telemedicine and Healthcare*. Final Report, December 2016.

[141] Al riguardo, v. le richiamate "*Guidelines on the protection of individuals with regard to the processing of personal data in a world of Big Data*" del Consiglio d'Europa del 23.1.2017, che raccomandano di considerare con attenzione non soltanto le implicazioni giuridiche, ma anche quelle sociali etiche e tecnologiche nell'uso dei dati (cfr., in particolare, punto 2).

[142] Cfr. Council of Europe, Consultative committee of the Convention for the protection of individuals with regard to automatic processing of personal data, *Practical guide on the use of personal data in the police sector*, Strasbourg, 15 February 2018, T-PD(2018)01.

Una conoscenza più approfondita del fenomeno – anzitutto da parte delle autorità di protezione dei dati personali, sia individualmente che nelle forme della rafforzata cooperazione ora introdotta con il RGPD – è necessaria anche perché le applicazioni delle tecniche *Big Data* nei vari settori della vita quotidiana richiederà i necessari aggiustamenti e, con ogni probabilità, l'interlocuzione con altri soggetti istituzionali, ad iniziare dalle autorità indipendenti di settore cui sono rimessi poteri di vigilanza e regolatori (si pensi già solo ai settori assicurativo, bancario, finanziario, energetico ecc.). il Garante è pienamente disponibile ad ogni forma di collaborazione sia ritenuta necessaria per conoscere e governare il fenomeno.

Se l'Indagine ha evidenziato la necessità che, anche presso le autorità di controllo, profili professionali (cd. *data scientist*) possano operare nel contesto dei *Big Data*, per assicurare anzitutto la qualità dell'attività di ricerca svolta tramite gli stessi, è del pari evidente che le competenze di tali figure professionali non possano prescindere da un'adeguata considerazione dei profili etici e giuridici (anzitutto con riguardo alle discipline di protezione dei dati personali) che tali trattamenti implicano.

Un'attenta ponderazione dei principi di protezione dei dati personali e l'impiego delle misure volte a prevenire la violazione dei diritti fondamentali degli interessati rappresentano quindi – seguendo l'invito del Consiglio d'Europa, per il quale "*preventive policies and risk-assessment shall consider the legal, social and ethical impact of the use of Big Data, including with regard to the right to equal treatment and to non-discrimination*"[143] – una cautela imprescindibile per lo sviluppo di tali tecnologie nel rispetto della dimensione individuale e collettiva dei diritti fondamentali (non limitati a quello alla protezione dei dati personali) in gioco.

Entro questa prospettiva valoriale si muoverà, nel contesto nazionale ed europeo, il Garante.

5. I *Big Data* nell'ecosistema digitale italiano: considerazioni dell'AGCM

5.1. *Big Data*, struttura di mercato e barriere all'entrata

Mercati e utilizzo dei *Big Data*. L'utilizzo dei *Big Data* è diffuso in una varietà di settori economici. In considerazione del grado di rilevanza dei *Big Data* nei processi competitivi, si possono distinguere almeno tre macro-categorie di settori/mercati, che risultano comunque in continua evoluzione:

i) mercati in cui l'utilizzo dei *Big Data* ha un rilievo minimo nella fornitura del bene/servizio. Si tratta di mercati nei quali i *Big Data* sono assimilabili ad altri *input* utilizzati dalle imprese, ad esempio, per migliorare la propria efficienza produttiva, senza però incidere in maniera significativa sul processo competitivo[144]. I *Big Data* possono essere impiegati dalle imprese per ottimizzare i processi interni di natura organizzativa e gestionale, per migliorare l'efficienza e la performance aziendale (ad esempio, per lo sviluppo di campagne di *marketing*, la gestione dei *call center* e la minimizzazione delle attività di frode, il c.d. *tracking* delle modalità d'uso e dello stato di salute dei macchinari, l'ottimizzazione della logistica e della distribuzione e la valutazione delle *performance* del personale);

[143] Council of Europe, Consultative Committee of the Convention for the Protection of Individuals with Regard to Automatic Processing of Personal Data (T-PD), *Guidelines on the protection of individuals with regard to the processing of personal data in a world of Big Data*, Strasbourg, 23 January 2017 T-PD(2017)01, p. 3.
[144] Cfr., al riguardo, audizione di Amazon (26 novembre 2018).

ii) mercati in cui l'utilizzo dei *Big Data* può incidere sulle condizioni di offerta del servizio, ad esempio in termini di qualità, e investe in maniera diretta la relazione fornitore-utente. Ciò può verificarsi nei settori caratterizzati da elevate asimmetrie informative - si pensi ai mercati "tradizionali" già *data intensive*, quali quelli finanziari, bancari e assicurativi - e dallo svolgimento di attività di distribuzione/intermediazione. In questi casi, la raccolta e l'analisi di una quantità sempre maggiore di dati porta a una conoscenza più approfondita dei processi e dei clienti e consente di adottare decisioni in grado di migliorare ogni aspetto dell'attività di impresa, dal *design* di prodotti e servizi, al *marketing*, alla vendita, al *customer care*;

iii) mercati in cui l'utilizzo dei *Big Data* è essenziale perché da esso dipendono caratteristiche fondamentali del bene/servizio, in particolare in termini di innovazione e/o di personalizzazione del servizio. Si tratta, ad esempio, dei servizi che rivestono un ruolo centrale nell'ecosistema digitale e che non potrebbero essere realizzati senza l'utilizzo di *Big Data*, ovvero di servizi in cui i *Big Data* permettono una "personalizzazione" dell'offerta, basata sulle caratteristiche del singolo utente. È evidente, ad esempio, che la caratteristica distintiva della pubblicità *online* risiede proprio nella capacità di utilizzare le informazioni raccolte sui singoli utenti per consentire agli inserzionisti pubblicitari di raggiungere *target* specifici di consumatori, indirizzando loro messaggi mirati, con crescenti livelli di personalizzazione (e di misurare in modo più preciso l'efficacia della campagna pubblicitaria).

A fronte della pervasività del fenomeno di *datafication* dell'economia, dunque, è soprattutto in alcuni ambiti che diviene impellente la necessità di considerare i *Big Data* nelle analisi economiche svolte per la comprensione del processo competitivo, anche nell'ambito dell'esercizio delle competenze di tutela della concorrenza dell'Autorità. A tal fine, occorre tenere presente che l'avvento e la diffusione di un'economia basata sui dati ha un impatto sia sulle dinamiche competitive che si realizzano all'interno di singoli mercati, sia sull'organizzazione e la fisionomia della catena del valore di interi settori economici.

Le caratteristiche economiche dei mercati *data-driven*. I modelli di *business* fondati sui *Big Data* costituiscono un aspetto che contraddistingue profondamente ecosistemi e servizi digitali, caratterizzati da elevati livelli di concentrazione e la presenza di operatori che detengono posizioni dominanti[145].

La disponibilità di *Big Data*, tuttavia, è solo *uno* dei diversi fattori che contribuiscono cumulativamente all'elevato grado di concentrazione e all'esistenza di barriere all'entrata nei mercati digitali. Infatti, altri fattori (oltre agli investimenti per sviluppare le capacità di analisi ed elaborazione dei dati), come le economie di scala e di scopo e le esternalità di rete, continuano a svolgere un ruolo importante nello spiegare il potere di mercato. Si tratta di aspetti che, pur non essendo nuovi nell'ambito dell'analisi *antitrust*, acquisiscono un particolare rilievo nei mercati digitali, per il

[145] In merito alle implicazioni della diffusione dell'economia digitale sulla definizione dei mercati rilevanti e sull'accertamento del potere di mercato si rinvia, da ultimo, a OCSE (2015), *Data Driven Innovation: Big Data for Growth and Well Being*, Paris; Cremer, J., de Montjoye Y.A., Schweitzer, H. (2019), *Competition Policy for the Digital Era*, Rapporto finale per la Commissione Europea; Furman, J. et al. (2019), *Unlocking digital competition, Report of the Digital Competition Expert Panel*, http://www.gov.uk/government/publications; George J. Stigler Center for the Study of the Economy and the State (2019), *Report of Committee for the Study of Digital Platforms, Market Structure and Antitrust Subcommittee*, The University of Chicago Booth School of Business.

condizionamento significativo che il loro effetto cumulato è in grado di esercitare sulle dinamiche concorrenziali.

In primo luogo, l'utilizzo dei *Big Data* tende spesso ad assumere rilievo in una specifica struttura di mercato, quella c.d. a **due o più versanti**, che si caratterizza per la presenza di due o più gruppi distinti di utenti e di effetti di rete diretti e/o indiretti[146]. Si tratta, ad esempio, di piattaforme *online* che danno luogo a mercati a più versanti c.d. di attenzione (come i motori di ricerca *online* o i *social network*[147]) o di scambio (es. i *marketplace* del commercio elettronico[148]).

In tale contesto, gli **effetti di rete** assumono particolare rilievo. Nelle c.d. piattaforme di attenzione, ad esempio, chi ha più utenti dispone di più dati per migliorare il proprio servizio, attirando a sua volta ancora più utenti e determinando così effetti di rete diretti, che si traducono in barriere all'uscita per gli utenti e in un più difficile ingresso per nuovi operatori[149]. Più dati consentono inoltre una maggiore capacità di generare valore nei confronti degli inserzionisti pubblicitari, aumentando i ricavi che potranno a loro volta essere investiti nella qualità del servizio, dando luogo a effetti di rete indiretti[150]. Tali effetti di rete possono progressivamente portare l'intero mercato a determinarsi a favore di una determinata piattaforma (c.d. *market tipping*), consolidandone la posizione[151].

In aggiunta agli effetti di rete, anche la presenza di significative **economie di scala e di scopo** può concorrere a determinare l'assetto dei mercati. Nel settore digitale, infatti, elevati costi fissi – che spesso possono risultare eccezionalmente elevati e non recuperabili (c.d "*sunk cost*"), come nel caso dei motori di ricerca[152] e delle piattaforme di distribuzione di *e-book*[153] – sono accompagnati da ridotti, o addirittura nulli, costi variabili.

Un ulteriore elemento che può incidere significativamente sul processo competitivo è dato dall'eventuale presenza di c.d. ***switching costs***, ovvero di limitazioni tecniche e/o economiche che possano prevenire gli utilizzatori dalla possibilità di cambiare fornitore. Ciò può avvenire, ad esempio, per assenza di interoperabilità tra sistemi di operatori concorrenti, generando fenomeni di c.d. *lock-in*, o per la scarsa propensione degli utenti a cambiare fornitore a causa dell'esistenza di significativi effetti di rete[154]. Per alcuni servizi digitali, il *multi-homing* può essere diffuso e contribuire a ridurre gli *switching costs* derivanti dagli effetti di rete[155]. Tuttavia, in alcuni casi, tale

[146] Le esternalità dirette di rete sussistono quando il beneficio che a un consumatore deriva dall'acquisto di un bene o di un servizio aumenta con il numero dei consumatori che acquistano lo stesso bene o servizio. Le esternalità indirette di rete derivano, invece, dal fatto che esistono due separati gruppi di utenti nei mercati a due versanti. Tali gruppi di utenti non beneficiano necessariamente solo dall'aumento della numerosità del proprio gruppo ma piuttosto, indirettamente, da quella dell'altro gruppo.

[147] Tali piattaforme normalmente forniscono un servizio gratuito a fronte di un sussidio derivante dalla pubblicità, alimentata dai dati e dall'attenzione degli utenti.

[148] Su tali piattaforme si realizzano di norma delle vere e proprie transazioni di scambio di beni o servizi.

[149] Cfr. audizione dei Proff. Preta, Maggiolino e altri (30 gennaio 2018); audizione del Prof. De Streel (19 febbraio 2018); audizione del dott. Quintarelli (13 settembre 2018).

[150] OCSE (2016), "*Big Data: bringing competition policy to the digital era - Background note by the Secretariat*", https://one.oecd.org/document/DAF/COMP(2016)14/en/pdf.

[151] Commissione Europea (2016), "*M.8124 Microsoft/Linkedin*", http://ec.europa.eu/competition/mergers/cases/decisions/m8124_1349_5.pdf, par. 340.

[152] Commissione Europea (2017), "*AT.39740 Google Search (Shopping)*", http://ec.europa.eu/competition/antitrust/cases/dec_docs/39740/39740_14996_3.pdf, par. 291.

[153] Commissione Europea (2017), "AT.40153 E-book MFNs and related matters (Amazon)", http://ec.europa.eu/competition/antitrust/cases/dec_docs/40153/40153_4392_3.pdf, par. 65.3.

[154] Cfr. audizione del Prof. Gambaro (18 dicembre 2017).

[155] Commissione Europea (2014), "*M.7217 Facebook/Whatsapp*",

fenomeno può essere limitato dai costi (anche in termini di costo opportunità) che gli utenti dovrebbero eventualmente sostenere per utilizzare attivamente una pluralità di piattaforme; ad esempio, il *multi-homing* potrebbe essere poco frequente nei servizi di *social network*, in considerazione del significativo tempo necessario a curare il proprio profilo sul singolo *network*[156].

Dati come barriera all'entrata. Nei mercati in cui l'utilizzo di *Big Data* assume un particolare rilievo nell'offerta del servizio e, dunque, nel processo competitivo, i *Big Data* possono costituire o contribuire a rafforzare le barriere all'entrata derivanti da altri fattori come le esternalità di rete.

Ad esempio, nei mercati a due versanti, l'accesso a una grande quantità di dati degli utenti può alimentare le esternalità di rete e costituire un significativo vantaggio competitivo, in quanto, per un potenziale nuovo entrante, potrebbe risultare difficile acquisire la quantità e la tipologia dei dati necessari per fornire un servizio adeguato ai propri clienti, nonché sviluppare la capacità necessaria per analizzarli. La combinazione tra utilizzo di *Big Data* ed effetti di rete può, dunque, consentire ai primi operatori che entrano sul mercato (c.d. *first movers*) di beneficiare di un significativo vantaggio competitivo rispetto ai potenziali nuovi entranti, creando barriere all'entrata.

Al fine di comprendere se e in che misura i *Big Data* costituiscano concretamente una barriera all'entrata per la fornitura di un particolare servizio, occorre tuttavia affrontare la questione in riferimento a uno specifico mercato e tenere presente (almeno) tre aspetti: *i)* la rilevanza dei *Big Data* per la fornitura del bene/servizio alla luce di tutte le caratteristiche del mercato in questione; *ii)* la natura, la qualità e la quantità di dati necessari per poter competere efficacemente; *iii)* la numerosità/varietà di fonti (sia *online* che *offline*) utilizzabili per generare la conoscenza rilevante per offrire i servizi in questione in maniera competitiva.

In primo luogo, dunque, rilevano le modalità concrete con le quali i *Big Data* sono utilizzati per la fornitura dello specifico servizio considerato, alla luce di tutte le caratteristiche del mercato interessato quali, ad esempio, la presenza e la natura delle esternalità di rete.

In linea generale, più che l'acquisizione dei dati in quanto tale, ciò che rileva a fini competitivi è dato dalle informazioni e dalla conoscenza generate attraverso i *Big Data*[157]. In alcune circostanze, a fronte dell'esistenza di algoritmi e meccanismi di *machine learning* consolidati, può essere la disponibilità di dati a costituire l'elemento effettivamente in grado di determinare un significativo vantaggio competitivo. In particolare, la disponibilità di grandi quantità di dati variegati per tipologia e provenienza può assumere rilievo cruciale ai fini del miglioramento della qualità degli algoritmi e della generazione di nuova conoscenza[158]. Può anche accadere, per converso, che i dati grezzi siano ampiamente disponibili e accessibili a tutti, ma è lo sviluppo di particolari algoritmi proprietari, attraverso investimenti ed innovazione, ad essere fonte di vantaggio competitivo[159].

---

http://ec.europa.eu/competition/mergers/cases/decisions/m7217_20141003_20310_3962132_EN.pdf, par. 135.

[156] Commissione Europea (2016), "*M.8124 Microsoft/Linkedin*", http://ec.europa.eu/competition/mergers/cases/decisions/m8124_1349_5.pdf, par. 345.

[157] Cfr. audizione dei Proff. Preta, Maggiolino e altri (30 gennaio 2018). "*L'idea alla base dei big data è l'utilizzo dei dati generati da un'attività per fini nuovi e diversi*" (audizione dei Proff. Giannotti e Pedreschi del 5 dicembre 2017).

[158] Cfr. audizione dei Proff. Giannotti e Pedreschi (5 dicembre 2017).

[159] A riguardo si veda, ancora, Cremer, J., de Montjoye Y.A., Schweitzer, H. (2019), *Competition Policy for the Digital Era*, Rapporto finale per la Commissione Europea; Furman, J. et al. (2019), *Unlocking digital competition, Report of the Digital Competition Expert Panel*, http://www.gov.uk/government/publications; George J. Stigler Center for the Study of

Un esempio dell'importante ruolo svolto sia dai dati che dagli algoritmi si rinviene nell'agricoltura digitale, che ha ad oggetto la raccolta di dati e informazioni sulle aziende agricole allo scopo di fornire a queste ultime servizi personalizzati o aggregati (come servizi di consulenza sui metodi di coltivazione). Nella valutazione dell'operazione di concentrazione *Bayer-Monsanto*[160], la capacità di fornire prescrizioni agronomiche è stata considerata dalla Commissione europea di fondamentale importanza per la competizione nel settore dell'agricoltura digitale. Essa richiede una serie di funzionalità rilevanti, quali la raccolta di dati agronomici, l'adozione di modelli automatici basati su algoritmi, nonché un sistema di consegna digitale di tali prescrizioni agli agricoltori tramite applicazioni o piattaforme[161]. In tale settore, la disponibilità di dati in quanto tale non è da sola sufficiente per poter fornire prescrizioni agronomiche digitali: occorre infatti elaborare un processo per "ripulire" i dati grezzi, un algoritmo per modellare e prevedere un dato fenomeno e un *software* in grado di combinare i dati e fornire quindi un servizio maggiormente competitivo.

In secondo luogo, rileva la natura, la qualità e la quantità di dati necessari per poter competere efficacemente. A tale riguardo, occorre tenere presente se e in che misura vi siano rendimenti di scala decrescenti (o crescenti) rispetto alla quantità di dati disponibili nonché rispetto alla varietà dei dati e alla velocità con la quale i dati sono raccolti[162]. Ad esempio, nel caso dei motori di ricerca *online*, che utilizzano dati per migliorare i risultati forniti agli utenti, maggiore è il numero di richieste che il motore di ricerca riceve, più rapidamente questo è in grado di cogliere un cambiamento nei comportamenti degli utilizzatori e di aggiornare e migliorare la rilevanza e la qualità delle proprie risposte. Tuttavia, i ritorni di scala in termini di miglioramento nella pertinenza dei risultati potrebbero decrescere una volta che il volume delle richieste ricevute superi un certo livello[163]. Laddove, invece, i dati vengono utilizzati prevalentemente per scopi pubblicitari e di *marketing* – stando a quanto emerso nel corso dell'Indagine – più che lo *stock* e il volume dei dati sembrerebbe che a rilevare sia la loro attualità. In questa prospettiva, le barriere per i nuovi operatori risulterebbero meno elevate, in quanto i dati dell'*incumbent* (e il suo vantaggio competitivo) sarebbero più rapidamente deperibili.

Infine, l'impatto dei *Big Data* sulle condizioni di concorrenza dipende dalla numerosità/varietà di fonti (sia *online* che *offline*) utilizzabili per generare la conoscenza rilevante finalizzata ad offrire i servizi in questione in maniera competitiva. Sebbene non si possa escludere a priori che alcuni *dataset* abbiano caratteristiche di unicità, va considerato che spesso le fonti utilizzabili per l'acquisizione delle informazioni rilevanti ai fini dell'offerta di un particolare bene/servizio sono molteplici. La

---

the Economy and the State (2019), *Report of Committee for the Study of Digital Platforms, Market Structure and Antitrust Subcommittee*, The University of Chicago Booth School of Business.

[160] Commissione Europea (2018), "*M.8084 Bayer/Monsanto*", http://ec.europa.eu/competition/mergers/cases/decisions/m8084_13335_3.pdf.

[161] I dati agronomici possono provenire dalle imprese (dati proprietari), da fonti pubbliche e di terze parti (es. dati meteorologici da satelliti o stazioni meteorologiche), dall'agricoltore mediante la fornitura manuale di dati (come il tipo di coltura o la varietà di semi) o da sensori collocati nei terreni dell'agricoltore. Mentre alcuni di questi dati sono relativamente facili da raccogliere (si pensi ad esempio alle immagini meteorologiche o satellitari) altri, come i dati proprietari e/o i dati degli agricoltori, sono nelle mani di un numero limitato di operatori. I dati agronomici proprietari vengono raccolti e accumulati attraverso attività di ricerca e sviluppo, verifiche sul campo, indagini di mercato.

[162] Cfr. Commissione Europea (2018), "*M.8788 Apple/Shazam*", http://ec.europa.eu/competition/mergers/cases/decisions/m8788_1279_3.pdf, par. 317.
Cfr. audizione dei Proff. Preta, Maggiolino e altri (30 gennaio 2018).
Cfr. audizione di IBM (22 ottobre 2018).

[163] Commissione Europea (2017), "*AT.39740 Google Search (Shopping)*", http://ec.europa.eu/competition/antitrust/cases/dec_docs/39740/39740_14996_3.pdf, par. 287.

maggior parte dei *dataset*, infatti, può essere in qualche modo replicata perché sui mercati sono disponibili numerosissimi dati, sia *online* che *offline*, molti dei quali pubblicamente accessibili e utilizzabili in modi alternativi[164]. Ad esempio, nell'ambito della valutazione dell'operazione di concentrazione *Facebook/WhatsApp*[165], la Commissione europea ha osservato come, anche a valle dell'operazione, vi sarebbe stata una elevata quantità di dati rilevanti a fini pubblicitari disponibile sul mercato. Anche nella valutazione dell'operazione *Apple/Shazam*[166], la Commissione ha considerato che l'integrazione dei *database* delle parti, contenenti dati sui rispettivi utenti, non avrebbe conferito alla nuova entità un vantaggio non replicabile. Si trattava, infatti, di database contenenti dati non unici e non qualificabili come input importanti per la fornitura dei prodotti a valle.

In ogni caso, solo in circostanze eccezionali, i *Big Data* raccolti da un'impresa possono costituire una risorsa "essenziale" per operare in un mercato, soprattutto in considerazione del fatto che la conoscenza necessaria per offrire i servizi può essere sviluppata attraverso l'impiego di *database* diversi (cfr. *infra*, sezione 5.4.4.).

Relazione tra diritto alla protezione dei dati personali e barriere all'entrata. Quanto rilevato in materia di *Big Data* e barriere all'entrata suggerisce che, in generale, esiste una possibile tensione tra diritto alla protezione dei dati e concorrenza. Poiché l'esistenza o meno di barriere all'entrata dipende anche dalla disponibilità di una molteplicità e varietà di fonti di acquisizione dei dati rilevanti, un approccio molto restrittivo volto a salvaguardare la protezione dei dati personali potrebbe avere come conseguenza un aumento delle barriere all'entrata. D'altro canto, va considerato che anche nell'esperienza nazionale – tenendo conto del quadro regolatorio proprio di singoli settori (ad es., energia elettrica e gas, servizi finanziari) – sono state individuate le condizioni, nel rispetto della disciplina di protezione dei dati, per consentire l'accesso a *dataset* di informazioni (selezionate) aventi natura personale riferite alla clientela, definendone le modalità di impiego[167]. In questa prospettiva, la piena applicabilità del RGPD potrebbe contribuire ad assicurare un livello di tutela omogeneo e favorire la costruzione di quel *level playing field* propedeutico allo svolgimento del processo competitivo.

Il portato concorrenziale della regolazione in materia di dati personali deve essere oggetto di particolare attenzione soprattutto laddove abbia un effetto differenziale, da una parte avvantaggiando le imprese *incumbent*, che tipicamente dispongono dei dati ottenuti dalla relazione diretta con i propri utenti, e dell'altra andando a svantaggio dei potenziali nuovi entranti, che potrebbero avere l'esigenza di acquisire con altre modalità i dati rilevanti per entrare e crescere nel mercato. Nella Sezione 5.3. di questo documento, la relazione non univoca tra concorrenza e protezione dei dati personali viene illustrata più in dettaglio.

Portabilità dei dati e riduzione *switching costs*. La portabilità dei dati può costituire un elemento di fondamentale rilevanza sotto il profilo concorrenziale. Riducendo i costi di *switching* dell'utente da una piattaforma all'altra, la portabilità dei dati può incidere, infatti, sulla mobilità degli utenti. La

[164] Cfr. audizioni di Microsoft (9 ottobre 2018), Facebook (5 febbraio 2018) e Amazon (26 novembre 2018).
[165] Commissione Europea (2014), "*M.7217 Facebook/Whatsapp*", http://ec.europa.eu/competition/mergers/cases/decisions/m7217_20141003_20310_3962132_EN.pdf, par. 188/189.
[166] Commissione Europea (2018), "*M.8788 Apple/Shazam*", http://ec.europa.eu/competition/mergers/cases/decisions/m8788_1279_3.pdf,.
[167] Cfr., a tale riguardo, il provv. del Garante 25 luglio 2007, n. 39, in http://www.gpdp.it, doc. web n. 1428567 e in G.U. n. 192 del 20 agosto 2007.

circolazione dei dati e la riduzione dei costi di *switching* possono contribuire a far sì che i dati non costituiscano una barriera all'ingresso, riducendo possibili rischi di *lock-in*, e che la mobilità degli utenti riduca gli effetti di rete connaturati all'attività delle piattaforme[168].

In questo contesto, il diritto alla portabilità dei dati, introdotto dal RGPD (cfr. art. 20), rappresenta senz'altro un importante passo avanti nella prospettiva di facilitare la circolazione dei dati e la mobilità degli utenti[169]. Si tratta infatti del diritto dell'interessato a ottenere i propri dati personali da un titolare del trattamento, qualora ricorrano talune condizioni quali il consenso dell'interessato o un contratto di cui l'interessato è parte (ai sensi dell'art. 6, par. 1, lett. a e b) e del diritto di trasmetterli ad altro titolare, senza impedimenti e in un formato strutturato, di uso comune e leggibile da dispositivo automatico[170].

Tuttavia, i possibili esiti dell'introduzione del diritto alla portabilità dipendono in maniera sostanziale dalla misura in cui tale diritto sarà esercitato dagli utenti. In quest'ottica, non può non rilevare come dal documento redatto dall'AGCM e pubblicato nel giugno 2018 di cui in Premessa emerga che attualmente solo 1 utente su 10 è consapevole dei propri diritti in materia di portabilità dei dati. Lo scarso interesse all'utilizzo della portabilità è dovuto alla bassa propensione ad utilizzare altre piattaforme/applicazioni (41,1%), ad una limitata sensibilità sulla rilevanza di tali dati (36,1%) nonché alla percezione di un'elevata complessità degli strumenti tecnologici (30,4%)[171].

5.2. Posizioni dominanti e potere di mercato

L'emergere di posizioni dominanti. A causa dell'effetto cumulato, e per certi versi esasperato, della struttura dei costi, degli effetti di rete diretti e indiretti, degli *switching costs*, della scarsa diffusione del *multi-homing* e dell'importanza crescente dei dati, i mercati digitali tendono ad essere concentrati e con elevate barriere all'entrata, determinando anche esiti di tipo c.d. "*winners take all*". L'acquisizione di una posizione di rilievo in un determinato mercato può essere ovviamente la conseguenza di una maggiore produttività o innovatività del prodotto o servizio offerto (e spesso lo è), alimentata dalle caratteristiche strutturali dei mercati digitali sopra descritte.

Tuttavia, alcuni servizi digitali che svolgono un ruolo centrale nell'ecosistema di Internet sono oggi controllati da operatori dominanti o che comunque non appaiono soggetti a pressioni competitive significative. Basti pensare alla posizione di Google nei servizi di ricerca *online* e nei sistemi operativi per dispositivi mobili, alla posizione di Facebook nei *social network* e a quella di Amazon nell'intermediazione nel commercio elettronico.

Il potere di mercato che i c.d. GAFA(M)[172] (Google, Apple, Facebook, Amazon e Microsoft) hanno assunto riveste peraltro rilevanza sistemica non solo per la dimensione globale dello stesso, ma anche per la circostanza che i servizi in questione hanno un ruolo centrale nell'abilitazione delle interazioni e transazioni digitali. Il controllo di tali *gateways*, dunque, consente di esercitare un'influenza significativa sulle dinamiche economiche e sociali che hanno luogo su Internet, gestendo di fatto

[168] Cfr. audizione dei Proff. Preta, Maggiolino e altri (30 gennaio 2018).
[169] Al riguardo si rinvia ad Article 29 Data Protection Working Party, Guidelines on the right to "*data portability*" (wp242rev.01),27/10/2017, adopted on 13 December 2016 as last revised and adopted on 5 April 2017, WP 242 rev.01.
[170] I dati devono essere stati oggetto di trattamenti automatizzati, con il consenso dell'interessato o sulla base di un contratto.
[171] Cfr. AGCM (2018), IC53 - Primi risultati dell'indagine conoscitiva sui *Big Data*, cit.
[172] Da ultimo, Tim Wu (2018), "*The Curse of Bigness' Review: Revisiting the Gilded Age*", Columbia University Press.

l'accesso ai mercati, la visibilità e la reputazione delle imprese terze che operano nell'ecosistema digitale e le loro relazioni con i consumatori finali[173].

L'integrazione verticale e conglomerale. L'integrazione verticale e conglomerale che caratterizza i principali operatori digitali costituisce un elemento di particolare rilievo nella valutazione del potere che tali operatori detengono nei singoli mercati rilevanti in cui sono attivi, nella misura in cui amplifica la loro capacità di acquisire, elaborare e sfruttare i dati nella fornitura dei servizi a consumatori e imprese. Ad esempio, la capacità di combinare i dati sul comportamento digitale di uno stesso soggetto, acquisiti da una varietà di fonti (e nel rispetto del quadro giuridico di riferimento), può consentire una profilazione estremamente puntuale e ben più sofisticata di quella che può derivare dall'acquisizione di dati parcellizzati sui singoli comportamenti di consumo digitale degli utenti.

In tale prospettiva, assumono rilievo diversi indicatori:

Ampiezza della gamma dei servizi offerti ai consumatori (e imprese): i grandi operatori digitali offrono generalmente numerosi servizi, spesso complementari, in una varietà di mercati rilevanti. L'offerta spazia dai dispositivi *hardware* connessi, ai sistemi operativi, a una molteplicità di applicazioni e servizi *online*. Ciò consente a tali operatori di intercettare i dati relativi ai comportamenti e alle abitudini di consumo degli utenti, pur rispettando il principio di finalità dei dati (che devono essere raccolti per finalità determinate, esplicite e legittime). Gli utenti, peraltro, sono spesso legati a un particolare "ecosistema", anche per i vantaggi derivanti dall'integrazione tra i diversi servizi che lo compongono e/o dalla limitata interoperabilità con i servizi offerti da operatori concorrenti.

La natura strategica di taluni servizi offerti ai consumatori: nell'ecosistema dei *Big Data* alcuni servizi – quali, ad esempio, i sistemi operativi, i *social network*, i motori di ricerca – rivestono una rilevanza particolare, sia in termini di capacità di acquisizione dei dati degli utenti sia per l'influenza determinante che hanno per una varietà di transazioni economiche (e sociali). Si tratta, peraltro, di servizi per i quali in diversi casi è stata già accertata l'esistenza di posizioni dominanti dalle autorità *antitrust* europee[174].

L'offerta alle imprese di servizi per l'acquisizione, la gestione e l'elaborazione di dati: i grandi operatori digitali sono spesso attivi non solo nell'offerta di servizi agli utenti finali, ma anche nell'offerta di servizi alle imprese per l'acquisizione, la gestione e l'elaborazione dei dati[175].

Ad esempio, nella fase di acquisizione dei dati è ricorrente la prassi secondo cui i siti *web*, anziché effettuare in proprio il tracciamento e la profilazione dei loro utenti e la computazione delle statistiche di accesso, ricorrono ai sistemi resi disponibili, anche gratuitamente, dai grandi operatori del *web* (come Google Analytics e Facebook Analytics), i quali in cambio ottengono l'accesso ai dati raccolti

[173] Cfr. audizione del Prof. De Streel (19 febbraio 2018).
[174] Commissione Europea (2019), "*AT.40411 Google Search (AdSense)*", cfr. http://europa.eu/rapid/press-release_IP-19-1770_it.htm; Commissione Europea (2018), "*AT.40099 Google Android*", cfr. http://europa.eu/rapid/press-release_IP-18-4581_it.htm; Commissione Europea (2017), "*AT.39740 Google Search (Shopping)*", cfr. http://ec.europa.eu/competition/antitrust/cases/dec_docs/39740/39740_14996_3.pdf; Bundeskartellamt (2019) B6-22/16 *Facebook, Exploitative business terms pursuant to Section 19(1) GWB for inadequate data processing*, cfr. https://www.bundeskartellamt.de/SharedDocs/Entscheidung/EN/Fallberichte/Missbrauchsaufsicht/2019/B6-22-16.pdf?__blob=publicationFile&v=4.
[175] Cfr., ad es., le audizioni IBM (22 ottobre 2018) e Microsoft (9 ottobre 2018).

dai siti stessi[176]. Gli operatori in questo modo acquisiscono dati puntuali relativi all'utilizzo dei siti/*app*, potendo così disporre delle statistiche relative all'accesso alle pagine, alla provenienza degli utenti e ai contenuti visualizzati[177]. I dati così raccolti, al netto dei profili di legittimità di cui alla normativa sulla protezione dati[178], possono essere combinati con i dati individuali che lo stesso operatore acquisisce grazie all'offerta di altri servizi nell'ambito del suo ecosistema, fino ad arrivare all'identificazione esatta dell'utente.

Anche nelle fasi successive alla raccolta dei dati, dove vengono offerti servizi utilizzati dalle imprese per la memorizzazione, l'elaborazione, l'analisi e l'interpretazione dei *Big Data*, alcuni dei grandi operatori digitali detengono posizioni di mercato di assoluto rilievo.

Il grado di integrazione tra i servizi offerti e di interoperabilità con servizi terzi: un importante elemento costitutivo del potere di mercato può essere rintracciato anche nei legami tra diversi servizi complementari presenti nella filiera dei *Big Data*, nonché nel limitato grado di interoperabilità con servizi terzi.

Ad esempio, con riferimento agli *standard* tecnologici per le piattaforme *cloud*, sono emerse due principali soluzioni concorrenti: *Amazon Web Services Compatible Solutions*, offerta direttamente da Amazon o da aziende con interfacce di programmazione delle applicazioni (*Application Programming Interfaces*) compatibili, e *OpenStack*, un progetto *open source* supportato da aziende del settore come IBM. La scelta di servizi compatibili con l'uno o l'altro di questi *standard* ha implicazioni sia sulla disponibilità degli strumenti *software* utilizzabili, sia sulla possibilità di rivolgersi a fornitori alternativi (*lock-in*).

La persistenza del potere di mercato. L'elevata concentrazione che si riscontra nei mercati digitali appare tanto più problematica quanto più tende ad assumere il carattere di persistenza negli anni. Lo slogan "*competition is a click away*" sembra mostrare segni di usura.

Le continue attività di espansione e diversificazione delle attività dei grandi operatori digitali determina una tensione concorrenziale che appare ancora idonea a generare innovazione e valore per i consumatori. Tuttavia, alcune posizioni di dominanza nella fornitura di taluni servizi non appaiono suscettibili di essere significativamente erose nel breve termine, anche nella prospettiva dinamica di concorrenza *per* il mercato, che pure ha generato una certa contendibilità delle posizioni di *leadership* nella fase iniziale di evoluzione di alcuni servizi digitali[179].

Infatti, l'esistenza di ecosistemi complessi, le rilevanti esternalità di rete che caratterizzano talune piattaforme multi-versante e la disponibilità di informazioni dettagliate sul comportamento dei propri consumatori sono solo alcuni dei fattori che appaiono rendere il potere di mercato dei grandi operatori digitali non solo particolarmente elevato, ma anche verosimilmente persistente.

---

[176] Cfr., al riguardo, la menzionata sentenza della Corte di giustizia (Grande Sezione) 1° ottobre 2019, causa C- 673/17, *Bundesverband der Verbraucherzentralen und Verbraucherverbände – Verbraucherzentrale Bundesverband eV c. Planet49 GmbH* nella quale sono esplicitate le condizioni (con particolare riguardo al consenso dell'interessato), richiamata anche *infra*.

[177] S. Englehardt & A. Narayanan, 2016, *Online Tracking: A 1-million-site Measurement and Analysis*, ACM CCS. Per considerazioni in ordine alla contitolarità del trattamento che si viene così ad integrare v. Corte di giustizia, 5 giugno 2018, C-210/16, *Unabhängiges Landeszentrum für Datenschutz Schleswig-Holstein/Wirtschaftsakademie Schleswig-Holstein GmbH*.

[178] D.C. Schmidt, 2018, *Google Data Collection*, DCN.

[179] Cfr. audizione del Prof. Gambaro (18 dicembre 2017).

Il rischio paventato da alcuni osservatori è che tali posizioni dominanti possano impedire in futuro l'entrata di nuovi operatori e ridurre gli incentivi all'innovazione e al miglioramento dell'offerta per gli *incumbent*, con effetti negativi sull'efficienza e il dinamismo delle imprese. Alcuni recenti studi, ad esempio, suggeriscono che, nei paesi OCSE e con particolare riguardo al settore digitale, a fronte di un margine di profitto medio conseguito dalle imprese in crescita, si assiste già ad una riduzione del tasso di entrata medio di nuove imprese nel mercato[180].

Il "nuovo" potere di mercato. Va osservato, inoltre, che la rivoluzione dell'economia digitale, oltre alle sfide di natura politica e tecnologica, solleva anche questioni di fondo sotto il profilo teorico, mettendo in discussione concetti di base come la stessa nozione di impresa, la definizione dei mercati o l'accertamento di potere di mercato[181]. Può sembrare paradossale - e ad oggi in parte ancora lo è - ma sempre più spesso (anche nel corso delle audizioni svolte nell'ambito dell'Indagine conoscitiva[182]) il "vero" potere di mercato tende ad essere attribuito a soggetti che non sono ancora presenti nel mercato e che, a detta degli attuali *incumbent*, in tempi ragionevolmente brevi sarebbero in grado di farvi ingresso, "distruggere" l'attuale assetto e raggiungere rapidamente una posizione difficilmente contestabile[183].

Una volta si sarebbero chiamati concorrenti potenziali molto aggressivi. E in parte lo sono, se si considera che, per questo timore, almeno alcuni degli operatori presenti nel mercato, anziché limitarsi a chiedere protezione, iniziano a sfruttare al meglio i dati settoriali e i vantaggi informativi derivanti dall'*incumbency*, che ancora oggi costituiscono vantaggi di posizione non irrilevanti[184]. In alternativa sarebbero definiti come imprese innovative in grado di "creare" nuovi mercati.

In entrambi i casi, tuttavia, la nuova sfida sotto il profilo *antitrust* può derivare dall'estrema difficoltà di fronteggiarli una volta che abbiano fatto ingresso nel mercato o ne abbiano creati di nuovi. Difficoltà che, secondo alcuni, originerebbe proprio dalla disponibilità di enormi volumi di dati e dalla capacità di analizzarli e elaborarli. Disponibilità e capacità che non sarebbero replicabili anche dai concorrenti "altrettanto efficienti".

È pertanto sostenibile che le grandi piattaforme digitali, i GAFA(M) o i c.d. *Datapolist*, come qualcuno ha iniziato a chiamarli[185], possono esercitare il loro potere ancor più che nei mercati dove sono già presenti (motori di ricerca, *social network*, *marketplace* e pubblicità *on-line),* nei mercati dove non sono ancora attivi ma in cui, grazie alla disponibilità di *Big Data* e alla capacità di elaborarli, potrebbero agevolmente entrare e rapidamente "dominarli", anche nel rispetto degli obblighi derivanti dalla disciplina in materia di protezione dei dati personali.

Nella misura in cui una simile ipotesi non sia del tutto infondata circostanza da misurare (anzitutto) alla luce del principio di finalità nel trattamento dei dati personali , la definizione del mercato

[180] Calligaris, S., C. Criscuolo and L. Marcolin (2018), "*Mark-ups in the digital era*", OECD Science, Technology and Industry Working Papers, 2018/10, OECD Publishing, Paris. http://dx.doi.org/10.1787/4efe2d25-en e Calvino, F. et al. (2018), "*A taxonomy of digital intensive sectors*", OECD Science, Technology and Industry Working Papers, 2018/14, OECD Publishing, Paris. http://dx.doi.org/10.1787/f404736a-en.

[181] Cfr. F. Jenny, Chair of OECD Competition Committee; OECD Competition Open Day, 27 February 2019, https://oecd.streamakaci.com/cod2019/.

[182] Cfr., tra le altre, audizioni Intesa Sanpaolo S.p.A. (23 febbraio 2018), UniCredit S.p.A. (8 marzo 2018).

[183] Cfr. audizione del dott. Quintarelli (13 settembre 2018) e audizione dei Proff. Preta, Maggiolino e altri (30 gennaio 2018).

[184] Cfr. audizione del Prof. Gambaro (18 dicembre 2017).

[185] Cfr. Ezrachi, A. e Stucke, E. (2016), *Virtual Competition: The Promises and Perils of the Algorithm Driven Economy*, Harvard University Press.

rilevante, strumentale all'accertamento del potere di mercato, va ripensata profondamente o, quanto meno, va dato uno spazio decisamente maggiore ai vincoli concorrenziali che possono essere sviluppati dal lato dell'offerta, tradizionalmente utilizzati per identificare la pressione della concorrenza potenziale: in altri termini, la pressione esercitabile dagli operatori presenti in mercati (merceologicamente e geograficamente) contigui.

Il fatto nuovo è che la disponibilità di *Big Data* sembrerebbe attribuire alle grandi piattaforme la capacità di esercitare una notevole disciplina concorrenziale su più mercati contemporaneamente, fino ad essere avvertite come soggetti dotati di un notevole potere di mercato ancor prima di avervi fatto ingresso.

L'impatto dei grandi operatori digitali, infatti, va anche al di là dei singoli mercati in cui operano, atteso che la forza *disruptive* dei processi di digitalizzazione può conformare l'organizzazione dell'intera filiera di settori digitali e tradizionali e ridefinire i rapporti negoziali che si instaurano tra gli operatori attivi in diverse fasi della filiera e/o in settori contigui.

Ad esempio, la digitalizzazione e le nuove modalità di distribuzione dei contenuti digitali hanno modificato radicalmente la filiera nel settore editoriale e giornalistico, la natura stessa del prodotto editoriale, i canali di distribuzione delle notizie, nonché i soggetti in grado di raccogliere gli investimenti pubblicitari all'interno di tale filiera e di appropriarsi dei contenuti[186].

In prospettiva, cambiamenti significativi, legati all'utilizzo dei dati e alla loro proprietà, potrebbero anche interessare i rapporti tra l'industria automobilistica e il settore assicurativo, anche in funzione del diverso grado di "controllo" che gli operatori avranno dei dati degli utenti. Ciò in uno scenario di mobilità connessa in cui assumerà rilievo anche un sistema digitalmente integrato che metta in connessione gli assicuratori e/o le case automobiliste, da un lato, e i clienti/conducenti, dall'altro, e che supporti servizi personalizzati e istantanei, riguardanti non solo il settore assicurativo, ad esempio, attraverso la definizione di polizze personalizzate, ma anche una serie di potenziali servizi attivabili direttamente dall'auto.

Il fermento registrato in entrambi i settori per sviluppare sistemi "integrati" è sintomo che il fattore temporale avrà un ruolo fondamentale nel determinare poi i rapporti di forza tra i due settori, soprattutto con riferimento alla proprietà e all'utilizzo dei *Big Data*, nell'offerta di servizi personalizzati che vanno anche al di là dell'ambito assicurativo. Come evidenziato dall'IVASS[187], il settore RC Auto italiano è ormai da alcuni anni soggetto a una evoluzione tecnologica, con un livello crescente di penetrazione di auto connesse attraverso scatole nere ovvero dispositivi elettronici che registrano l'attività dei veicoli sui quali sono installati offerti dalle compagnie assicurative o altri dispositivi elettronici offerti dalle case automobilistiche.

Inoltre, nel corso dell'indagine, gli operatori dei settori delle informazioni creditizie, bancario-assicurativo e delle telecomunicazioni, tradizionalmente caratterizzati da un quadro regolatorio strutturato e stringente, hanno manifestato l'esigenza, in maniera trasversale, che venga assicurato un *level playing field* al fine di consentire alle imprese "tradizionali" di competere con gli OTT senza il vincolo di asimmetrie regolatorie[188]. Tale istanza appare riconducibile alla possibilità che gli OTT

[186] Cfr. audizione di IlSole24Ore (1 dicembre 2017).
[187] https://www.ivass.it/consumatori/azioni-tutela/indagini-tematiche/documenti/2018/AnalisiTrend1sem2018.pdf.
[188] Cfr. audizioni di Unicredit (8 marzo 2018), Intesa San Paolo (23 febbraio 2018), Generali (21 marzo 2018), Experian (28 novembre 2017), CRIF (18 dicembre 2017), Allianz (17 novembre 2017), Vodafone (7 dicembre 2018), Wind-Tre (29 novembre 2018), Fastweb (7 dicembre 2018), TIM (7 dicembre 2018).

espandano la propria attività, basata sui dati, a settori tradizionali come quelli delle informazioni creditizie e bancario-assicurativo e, per altri versi, che le imprese di telecomunicazioni possano espandere la propria operatività a servizi basati sui dati.

Potere di mercato e operazioni di concentrazione. In alcuni casi, la creazione o il rafforzamento di potere di mercato in mercati *data-driven* derivano da fenomeni di crescita esterna.

Si tratta di operazioni che possono sfuggire al controllo delle concentrazioni previsto dalle norme a tutela della concorrenza e che riguardano principalmente acquisizioni da parte di operatori dominanti di *startup* potenzialmente *disruptive* (le cosiddette "*killer acquisitions*"), soprattutto quando oggetto dell'acquisizione sono i dati e la capacità di analizzarli. Il recente dibattito internazionale sull'adeguatezza di un controllo preventivo delle operazioni di concentrazione il cui ambito è definito, esclusivamente o quasi, da un sistema di notifica basato su soglie di fatturato e sulla opportunità di colmare eventuali *gap*, ha portato Germania ed Austria a introdurre criteri di notifica basati sul valore della transazione e il requisito che l'impresa oggetto di acquisizione sia attiva in maniera considerevole in tali Paesi.

Consentire un controllo delle concentrazioni efficace e rigoroso nei mercati digitali costituisce un obiettivo di *policy* che deve essere ampiamente condiviso, sia a livello nazionale che internazionale.

Nei mercati in cui i *Big Data* assumono rilievo, imprese e consumatori prendono decisioni sulla base di diversi aspetti dei prodotti/servizi considerati. Accanto al prezzo, che può anche essere nullo, assumono rilievo anche il grado innovazione e il livello di qualità dei prodotti/servizi offerti, anche con riguardo al livello di protezione dei dati offerto.

Con riferimento ai profili di merito delle analisi delle concentrazioni, dunque, occorre assicurare che l'analisi degli effetti delle operazioni di concentrazione sia adeguata a cogliere le peculiarità di mercati *zero-price*, in cui assumono rilievo centrale altre dimensioni del confronto competitivo quali il grado di innovazione e il livello di qualità dei servizi e della protezione dei dati degli utenti[189].

Per quanto riguarda il primo aspetto, nell'ambito della valutazione delle operazioni di concentrazione, approfondire l'impatto dell'innovazione e della qualità sulla concorrenza, oltre agli effetti sui prezzi, rappresenta una sfida importante e ciò anche per la relazione complessa che può caratterizzare il rapporto tra pressione concorrenziale e qualità. Introdurre più esplicitamente e analiticamente l'analisi di aspetti diversi dal prezzo monetario all'interno della valutazione degli effetti delle operazioni di concentrazione costituisce una sfida complessa.

Con riguardo all'innovazione, ad esempio, l'estesa letteratura economica evidenzia come il rapporto con la concorrenza dipende da una varietà di aspetti di non agevole considerazione a livello analitico[190], anche a causa delle necessità di dover fare valutazioni prospettiche in relazione a mercati che esibiscono comunque un notevole dinamismo. Peraltro è ben possibile che un'operazione di concentrazione possa determinare un aumento della capacità ad innovare, attraverso ad esempio una crescita della dimensione di impresa e della combinazione di attività complementari, anche a fronte di una riduzione degli incentivi derivanti dall'eventuale perdita di una pressione competitiva.

[189] Cfr. OCSE (2018), *Quality Considerations in Digital-Zero Price Markets*, *Background note by the Secretariat*, Parigi, 28 novembre.

[190] Ad esempio, occorre considerare la tipologia di attività innovativa svolta dalle parti, la struttura dei mercati del prodotto collegati a tale attività innovativa, le caratteristiche della concorrenza di natura dinamica, e la capacità delle imprese di appropriarsi dei benefici dell'innovazione.

Anche la qualità può rappresentare un elemento del benessere del consumatore rilevante quanto il prezzo e un aspetto centrale della strategia competitiva delle imprese, soprattutto nei mercati *zero-price*. Alcune operazioni di concentrazione possono comportare una riduzione unilaterale della qualità, esattamente come possono determinare un incremento dei prezzi, sebbene in un contesto in cui assume particolare rilievo il fenomeno della differenziazione strategica del prodotto[191]. Un'ulteriore criticità per le autorità di concorrenza emerge laddove sia necessario considerare l'interazione degli effetti di prezzo e non di prezzo. In particolare, quando gli effetti non di prezzo sono rilevanti e importanti, può essere difficile bilanciare tali effetti con eventuali effetti sui prezzi, in particolare se si muovono in direzioni diverse.

Con particolare riguardo ai mercati *data-driven,* la protezione dei dati individuali, la trasparenza e le informazioni necessarie per una scelta consapevole del consumatore possono rappresentare fattori qualitativi rilevanti per il confronto concorrenziale tra piattaforme digitali, soprattutto in mercati caratterizzati da prezzi nulli. I consumatori possono, infatti, preferire soluzioni che consentano di fornire la quantità minore di dati possibile o di mantenere il maggior controllo possibile sull'utilizzo dei dati personali forniti.

Qualche anno fa si registrava una certa ritrosia a introdurre esplicitamente considerazioni riguardanti la protezione dei dati nell'ambito della valutazione degli effetti delle operazioni di concentrazione. Ad esempio, nella concentrazione *Facebook/WhatApp*[192], autorizzata senza condizioni nel 2014, la Commissione europea ha considerato che la *privacy* non fosse un importante parametro nella scelta dei consumatori nel mercato delle applicazioni di comunicazione tra i consumatori e che, quindi, una possibile degradazione della *privacy* - a valle dell'operazione di concentrazione - non avrebbe potuto danneggiare il benessere degli stessi.

Negli ultimi anni, tuttavia, è emersa una tendenza diversa nella pratica delle autorità di concorrenza, fondata su una interpretazione ampia della nozione di benessere dei consumatori, tale da ricomprendere correttamente anche la dimensione della protezione dei dati personali. Ad esempio, nella recente concentrazione *Microsoft/Linkedin*[193], autorizzata con condizioni nel 2016, la Commissione ha concluso che la *privacy* è un importante parametro della concorrenza tra *social network* professionali e che, in assenza di rimedi adeguati, la concentrazione avrebbe potuto escludere concorrenti in grado di offrire una *privacy* maggiore di quella offerta da Linkedin, fermo restando il rispetto del quadro regolamentare in materia di protezione dei dati personali.

Introdurre considerazioni di *privacy* all'interno del controllo delle concentrazioni, tuttavia, costituisce solo uno dei diversi strumenti a disposizione dell'autorità di concorrenza per contribuire alla tutela di tale valore nell'esercizio delle proprie competenze. Come illustrato nella sezione dedicata al rapporto tra concorrenza e *privacy* (Sez. 5.3), infatti, una relazione virtuosa tra questi due aspetti può svilupparsi solo laddove i consumatori assumano scelte di consumo dei servizi digitali anche sulla base di una effettiva consapevolezza e sensibilità rispetto alla tutela dei propri dati.

[191] Cfr. Federico G. et al. (2017), "*A simple model of mergers and innovation*", Economic Letters, Vol. 157; Argentesi, E. et al. (2016), *The effect of retail mergers on prices and variety: an ex-post evaluation*, DICE Discussion Paper, 225.
[192] Commissione Europea (2014), "*M.7217 Facebook/Whatsapp*", http://ec.europa.eu/competition/mergers/cases/decisions/m7217_20141003_20310_3962132_EN.pdf.
[193] Commissione Europea (2016), "*M.8124 Microsoft/Linkedin*", http://ec.europa.eu/competition/mergers/cases/decisions/m8124_1349_5.pdf.

Un paio di considerazioni a parte richiedono le implicazioni per il controllo delle concentrazioni determinate dalla diffusione degli algoritmi di prezzo. Ancora oggi le evidenze in merito alla portata e ai pro e contro del fenomeno sono piuttosto fragili. Tuttavia, appare ragionevole ipotizzare che l'accresciuta possibilità di praticare strategie di prezzo discriminatorie richieda una maggiore sensibilità, sia sulla definizione del mercato rilevante del prodotto (che potrebbe inevitabilmente collassare su quelli che, in assenza di algoritmi di prezzo, erano segmenti di un mercato più ampio) sia sulla valutazione degli effetti unilaterali. Lo sviluppo di algoritmi di prezzo pro-collusivi comporta invece una maggiore attenzione nelle valutazioni dei rischi di coordinamento anche in mercati più frammentati[194].

Più in generale, inoltre, la crescente rilevanza assunta dai *Big Data* in alcuni settori suggerisce di guardare alle acquisizioni di natura conglomerale con un'attenzione maggiore di quella tradizionalmente loro riservata. La disponibilità di dati e la capacità di analizzarli può infatti consentire il rafforzamento del potere di mercato in mercati anche apparentemente "lontani" tra loro.

Infine, occorre considerare l'opportunità di introdurre a livello normativo criteri di notifica delle operazioni di concentrazione che consentano il controllo preventivo di operazioni di fusione e acquisizione di imprese innovative da parte dei grandi operatori digitali, le quali oggi possono sfuggire al vaglio dell'Autorità a causa della natura e del livello delle soglie previste per le comunicazioni.

5.3. *Big Data*, utilizzo dei dati personali e concorrenza

5.3.1. Premessa

Nel corso dell'Indagine è emerso come i dati siano sovente trattati per scopi definiti solo in termini generali: l'acquisizione massiva dei dati rende infatti difficoltosa la specifica individuazione *ex ante* delle finalità del relativo trattamento (cfr. *supra*, § 2.5).

Tuttavia, laddove si tratti di dati personali, il RGPD prevede che le attività di raccolta e utilizzazione dei dati possano avvenire previa richiesta del consenso dell'interessato o al ricorrere di una delle condizioni previste dall'art. 6. È inoltre stabilito (cfr. art. 5) che i dati personali siano trattati in modo lecito, corretto e trasparente, siano raccolti e trattati per finalità determinate, esplicite e legittime e siano adeguati, pertinenti e limitati a quanto necessario rispetto alle finalità per le quali sono trattati (principio di minimizzazione dei dati), siano esatti e se necessario aggiornati, nonché conservati in modo idoneo ad identificare gli interessati per il tempo necessario al conseguimento delle finalità per le quali sono trattati e siano trattati in modo da garantirne un'adeguata sicurezza.

Nello specifico, il principio di minimizzazione impone che, qualora il titolare intenda raccogliere dati ulteriori rispetto a quelli in suo possesso o trattare i dati per una finalità diversa rispetto a quella comunicata, dovrà richiedere il relativo consenso all'interessato. Questa operazione, apparentemente lineare, difficilmente si concilia con l'acquisizione di enormi quantitativi di dati, dovendosi concretamente determinare, volta per volta, l'utilizzo che dei dati si andrà ad effettuare, al fine di limitare la loro raccolta a quanto necessario per svolgere il servizio offerto. Come anche osservato

[194] Cfr. Harrington, J. (2018), *Developing Competition for Collusion By Autonomous Artificial Agents*, Working Paper, The Wharton School, University of Pennsylvania; McSweeney, T. e O Dea B., *The Implications of Algorithmic Pricing for Coordinated Effects Analysis and Price Discrimination Markets in Antitrust Enforcement*, Antitrust, Vol. 32,No. 1, Fall 2017.

dal Garante per la protezione dei dati personali[195], i principi di minimizzazione, limitazione della finalità e conservazione per il solo tempo indispensabile alla realizzazione del trattamento non si attagliano a raccolte così massive di dati, acquisiti spesso non per esigenze attuali ma in vista di future, eventuali necessità e riutilizzati per fini ulteriori non sempre compatibili con quelli originari.

Già il Gruppo di lavoro Articolo 29, nella dichiarazione del settembre 2014 relativa all'impatto dello sviluppo dei *Big Data* sulla protezione delle persone rispetto al trattamento automatizzato dei loro dati personali nell'UE, affermava che le caratteristiche intrinseche dei *Big Data* richiedono l'adozione di soluzioni innovative per far sì che i principi in materia di protezione dei dati personali possano concretamente applicarsi[196].

Nell'Indagine sono state proposte soluzioni innovative volte a favorire la partecipazione dell'individuo nel trattamento dei propri dati con tecniche di *Big Data,* come il ricorso al *dynamic consent*, sviluppato nel contesto delle c.d. bio-banche. Secondo questo modello, l'individuo presta inizialmente un consenso ampio a fronte di un'informativa generale circa le possibili finalità del trattamento e, successivamente (una volta individuata specificatamente la finalità di utilizzo dei dati), riceve una più puntuale informativa con la richiesta di un nuovo e più specifico consenso al trattamento (cfr. *supra*, § 2.5).

Sia l'applicazione della normativa sulla protezione dei dati personali che la strumentazione propria della tutela del consumatore possono offrire un contributo importante per la riduzione dell'asimmetria informativa esistente, garantendo che gli utenti ricevano un'adeguata informazione circa le finalità della raccolta e dell'utilizzo dei loro dati e siano posti nella condizione di esercitare consapevolmente ed effettivamente le proprie scelte di consumo.

La raccolta e soprattutto l'utilizzo di dati personali assumono interesse anche nella prospettiva del diritto della concorrenza nella misura in cui i dati si configurano come "beni economici" idonei a generare un profitto per le imprese.

Le autorità di concorrenza sono consapevoli del fatto che l'attività di profilazione – resa possibile dall'acquisizione massiva dei dati - portata ai suoi estremi, può agevolare comportamenti abusivi idonei a ridurre la contendibilità degli ecosistemi delle principali piattaforme, rendendo persistente il loro potere di mercato. Anche la diffusione di algoritmi di prezzo, anch'essa agevolata dalla disponibilità di grandi quantitativi di dati, può facilitare la stabilità di cartelli e la creazione di contesti di mercato favorevoli ad equilibri collusivi.

È dunque evidente che l'utilizzo dei *Big Data* interessi ambiti di competenza e di criticità di diversa natura e che le sfide poste dallo sviluppo dell'economia digitale richiedono l'impiego sinergico degli strumenti a tutela della *privacy*, del consumatore e della concorrenza.

---

[195] Cfr. *supra*, § 4.7., nonché intervento di Antonello Soro, Presidente del Garante per la protezione dei dati personali ("GARRNEWS", 23 agosto 2018), *Big Data e Libertà nella dimensione digitale*, *sub* file:///D:/Users/agcmcd/Downloads/GarantePrivacy-9036954-1.1.pdf.

[196] *Statement of the WP29 on the impact of the development of big data on the protection of individuals with regard to the processing of their personal data in the EU*, adottata il 16 settembre 2014. Cfr. altresì le *Guidelines on the protection of individuals with regard to the processing of personal data in a world of Big Data* del Consiglio d'Europa, adottate il 23 gennaio 2017, ove si legge che "*Given the transformative nature of the use of Big Data and in order to comply with the requirement of free, specific, informed and unambiguous consent and the principles of purpose limitation, fairness and transparency, controllers should also identify the potential impact on individuals of the different uses of data and inform data subjects about this impact*".

Fatte queste doverose premesse, nei paragrafi successivi saranno analizzate le numerose fonti di estrazione di valore economico dei dati (personali e non), che i mercati digitali hanno potuto sviluppare, e si tenterà di verificare l'impatto della disponibilità di tali dati sull'utilità del consumatore e sulle dinamiche concorrenziali. Resta tuttavia impregiudicata ogni valutazione circa la legittimità, dal punto di vista della normativa a tutela della *privacy*, delle attività di raccolta e utilizzo dei dati personali, la cui "necessarietà" e "pertinenza" dovrà essere verificata caso per caso in relazione alla finalità per la quale i dati sono stati acquisiti e il servizio offerto. L'eventuale eccedenza dei dati rispetto allo scopo per il quale sono stati raccolti o utilizzati potrà comunque rilevare anche sotto il profilo *antitrust* nella misura in cui sia idonea a comportare restrizioni concorrenziali nei mercati interessati[197].

5.3.2. L'acquisizione di dati personali nel processo produttivo e benessere del consumatore

Lo sviluppo dei mercati digitali non solo ha ampliato la quantità di dati personali che possono essere impiegati nel processo produttivo, ma ha fatto emergere nuove fonti di raccolta e forme di utilizzo dei medesimi.

Lo sviluppo dell'*information technology*, ed in particolare delle cosiddette tecnologie Web 2.0 (*blogs*, *social media*, *on line social networks*) ha posto al centro dell'attenzione l'individuo, non tanto come mero consumatore, quanto piuttosto come "produttore/generatore" di dati personali successivamente utilizzati nei processi produttivi di molte imprese operanti in una varietà di mercati[198].

Sotto il profilo dell'acquisizione, l'avvento della nuova era digitale ha, in primo luogo, amplificato la materiale disponibilità di dati personali pubblicamente accessibili (attraverso le informazioni che gli utenti rilasciano anche volontariamente sui *social network*, come Facebook, Instagram e Linkedin) e di quelli acquistabili sul mercato (ad esempio presso i cosiddetti *data broker*, soggetti che aggregano informazioni sui consumatori rinvenibili da diverse fonti pubbliche). Inoltre, le imprese in molti settori possono più facilmente validare i dati che acquisiscono nell'ambito del rapporto contrattuale con i propri utenti, ad esempio incrociando un proprio *database* clienti già esistente con altri *dataset*, legittimamente trattati, acquistabili sul mercato (si pensi ad una banca che è in grado di verificare le informazioni sul rischio di credito dei propri clienti con dati ricavabili da altre fonti informative).

L'impatto più dirompente, da valutare anche alla stregua dei principi di pertinenza e non eccedenza nonché di minimizzazione nell'utilizzo dei dati (contenuti nel RGPD), appare però rinvenirsi nell'opportunità per i fornitori di molti servizi *online* di acquisire informazioni sui propri utenti anche ulteriori rispetto a quelle strettamente connesse all'oggetto del contratto, sulla base del consenso dell'interessato al trattamento dei propri dati personali per una o più specifiche finalità[199]. In

[197] Per una discussione della relazione esistente tra protezione dei dati personali, tutela della concorrenza e tutela del consumatore si veda: Costa-Cabral, Francisco and Lynskey, Orla (2017) Family ties: the intersection between data protection and competition, *EU Law.Common Market Law Review*, 54 (1). pp. 11-50; EDPS Opinion on coherent enforcement of fundamental rights in the age of big data, 8/2016.

[198] Cfr. E. Posner & E. Glen Weyl, *Radical Markets: Uprooting Capitalism and Democracy for a Just Society*, 2018.

[199] Cfr., con riguardo al trattamento necessario all'esecuzione di un contratto di cui l'interessato è parte o all'esecuzione di misure precontrattuali adottate su richiesta dello stesso, il documento dell'EDPB, *Guidelines 2/2019 on the processing of personal data under Article 6(1)(b) GDPR in the context of the provision of online services to data subjects*, Version 2.0, 8 October 2019. Anche il Garante per la protezione dei dati personali ha ritenuto illecito il trattamento consistente nel conferimento obbligatorio di dati personali, in violazione del principio di minimizzazione oltre che di correttezza nel trattamento, "autorizzandone" l'uso per finalità di marketing, già solo per accedere ad un sito di e-commerce, con il provv. 20 luglio 2017, n. 324, doc. web n. 6955363.

particolare, gli sviluppatori di applicazioni su dispositivi mobili (ad esempio, *smartphone*) possono accedere, da un lato, ai dati che gli utenti immagazzinano su detti dispositivi, in genere sulla base del loro consenso, e dall'altro ad informazioni circa l'utilizzo di altre applicazioni installate che ne arricchiscono il profilo. Proprio per sfruttare tale opportunità di acquisizione dei dati, anche imprese operanti in mercati più tradizionali hanno iniziato a sviluppare applicazioni di interfaccia con i propri utenti. In altri termini, non solo sono cresciute le fonti di acquisizione, ma tale processo di raccolta dei dati tende a riguardare un numero sempre maggiore di operatori.

In concomitanza con lo sviluppo di nuove fonti di raccolta dei dati, sono sorte nuove opportunità di utilizzo degli stessi che ne hanno favorito il processo di estrazione del valore economico, con implicazioni potenzialmente diverse sul benessere dei consumatori e sul benessere sociale.

Prima dello sviluppo dei mercati digitali, l'acquisizione dei dati per le imprese era tendenzialmente volta a ridurre le asimmetrie informative riguardanti le caratteristiche dei propri clienti che, dando luogo a problemi di selezione avversa, potevano generare inefficienze (allocative) sul mercato. Tornando all'esempio della banca, una maggiore disponibilità di informazioni per valutare l'affidabilità del debitore consente di definire un tasso di interesse più basso, con l'effetto di aumentare la platea di clienti disponibili a pagare un prezzo per farsi finanziare i propri investimenti e di abbassare il rischio di credito verso questi ultimi. Pertanto, sotto questa prospettiva, fermo restando il rispetto dei principi di pertinenza/non eccedenza e minimizzazione, l'accesso ad un maggiore ammontare di dati personali può ridurre il rischio che risorse economiche e fattori produttivi (in questo caso di tipo finanziario) vengano impiegati in maniera inefficiente.

Accanto a questa modalità di impiego - che favorendo una migliore allocazione delle risorse non pone problemi in termini di benessere sociale e del consumatore - i mercati digitali hanno sviluppato nuove prospettive di utilizzo dei dati personali. In primo luogo, dal punto di vista tecnico (ed impregiudicata ogni valutazione giuridica sulla protezione dei dati), la possibilità di tracciare e raccogliere informazioni sull'utilizzo di dispositivi mobili o sull'attività di ricerca svolta via *web* dagli utenti consente alle piattaforme *online* di inferirne le preferenze, le abitudini e i potenziali bisogni di consumo[200]. Da queste informazioni è possibile ricavare una puntuale profilazione dell'utente, che può avere diverse finalità per le imprese, dal miglioramento dell'esperienza di fruizione del servizio all'offerta di nuovi servizi "contigui", o ancora ad una personalizzazione della comunicazione pubblicitaria, alla possibilità di praticare prezzi differenziati che riflettono la disponibilità a pagare di ciascuno (ma che pure potrebbero sottendere pratiche abusive)[201].

Un'altra fonte di estrazione di valore economico dai dati personali deriva dalla possibilità di un utilizzo secondario dei dati raccolti (da misurarsi anche con il principio di compatibilità con la finalità per la quale il consenso è stato prestato, ai sensi dell'art. 5, par. 1, lett. b) del RGPD), che non solo può travalicare l'oggetto del contratto, non risultando dunque complementare al servizio (primario) offerto ai clienti, ma può comportare una perdita totale del controllo da parte dei clienti sulle informazioni rilasciate al proprio fornitore del servizio. Tale circostanza si verifica quando quest'ultimo utilizza i dati personali raccolti (servizio secondario), per svolgere un'attività che si pone sull'altro versante del mercato, da cui il fornitore del servizio principale acquisisce un'ulteriore fonte di profitto. Anche in questo contesto, non è pacifico valutare se il consumatore sia danneggiato o

[200] Cfr. audizione del Prof. Giustozzi (16 novembre 2017) e audizione di Mediaset (28 novembre 2017).
[201] Cfr. audizione del Prof. Gambaro (18 dicembre 2017).

meno dal trasferimento di dati al secondo versante del mercato. Da un lato, infatti, il servizio primario è offerto "gratuitamente" dalle imprese, e dunque i profitti generati dalla vendita a terzi dei dati personali consentono di sussidiare l'erogazione del servizio principale. Dall'altro lato, i dati possono essere utilizzati per attività che incidono in vario modo sul benessere dei consumatori.

L'impatto che l'utilizzo dei dati ha sull'offerta di beni e servizi fa venire in rilievo diversi scenari in merito alla relazione tra *Big Data* e benessere dei consumatori:

i) *L'utilizzo dei dati personali riduce il benessere dei singoli consumatori.* Vi sono situazioni nelle quali l'utilizzo dei dati personali dell'utente può determinare una riduzione del benessere di quest'ultimo. Ad esempio, il tracciamento dell'attività di ricerca *online* dell'utente di prodotti o di servizi finalizzato ad un acquisto che poi non viene perfezionato può essere sfruttato per proporre al consumatore in un momento successivo lo stesso prodotto o servizio ad un prezzo maggiorato. In questo contesto, a fronte di un incremento del *surplus* delle imprese, si realizza un deterioramento delle condizioni economiche dei consumatori.

ii) *L'utilizzo dei dati personali aumenta il benessere dei singoli consumatori.* Vi sono diverse situazioni in cui l'utilizzo dei dati personali da parte delle imprese comporta un miglioramento del benessere dei consumatori in termini di qualità, varietà e condizioni economiche dei servizi disponibili.

a. *Innovazione.* In alcuni casi, la raccolta dei dati personali è necessaria per la fornitura di un certo servizio. Ad esempio, la raccolta e l'elaborazione dei dati sulla localizzazione degli utenti ha permesso lo sviluppo di servizi che forniscono in tempo reale informazioni sul traffico. Nella prospettiva del singolo consumatore sarebbe ottimale avere accesso alle informazioni in tempo reale sul traffico nelle città senza condividere i dati relativi alla propria localizzazione. Tuttavia, non sarebbe possibile realizzare il servizio in questione senza accesso ai dati relativi ai singoli utenti: la condivisione dei dati personali costituisce il presupposto stesso per lo sviluppo di un tale servizio.
b. *Qualità e varietà dei beni e servizi disponibili.* I dati acquisiti, inoltre, possono consentire la fornitura all'utente di un servizio di qualità migliore, come nel caso in cui le piattaforme digitali suggeriscono prodotti e servizi di interesse per il singolo consumatore, con un beneficio per quest'ultimo in termini di riduzione dei costi di ricerca e transazione nel mercato.

iii) *L'utilizzo dei dati personali aumenta il benessere dei singoli consumatori, ma riduce il benessere sociale.* Occorre osservare, infine, che vi possono essere situazioni particolari in cui l'utilizzo dei dati personali aumenta il benessere dei singoli consumatori, ma riduce il benessere sociale, a causa ad esempio dell'esistenza di esternalità negative. A titolo esemplificativo, la personalizzazione dei contenuti giornalistici proposti agli utenti dalle piattaforme di ricerca e *social network* può essere gradita dal singolo utente, che ha accesso a contenuti di interesse e in linea con le proprie preferenze, ma può risultare non desiderabile per la società nella misura in cui riduce il grado di pluralismo nel consumo dei contenuti giornalistici con un impatto sulla sfera politica e sociale. In queste situazioni, dunque, è rilevante distinguere l'obiettivo di *policy* della tutela del consumatore da obiettivi diversi quali la tutela del pluralismo nell'informazione.

In questa prospettiva si pongono i seri rischi, come hanno evidenziato casi recenti, di utilizzo improprio dei dati personali delle persone per sofisticate attività di profilazione su larga scala e di invio massivo di comunicazioni o campagne personalizzate (il c.d. *micro-targeting*) volte a influenzare l'orientamento politico e/o la scelta di voto degli interessati, sulla base degli interessi personali, dei valori, delle abitudini e dello stile di vita dei singoli. La corretta applicazione delle norme sulla protezione dei dati, soprattutto *on-line*, è essenziale strumento di protezione, ad es. dei processi elettorali, da interferenze e turbative esterne e, in ultima analisi, garanzia di corretto funzionamento della democrazia[202].

Un aspetto diverso è se e in che misura, pur nell'ambito dell'esistente quadro regolamentare, l'assetto delle interazioni tra utenti ed imprese in tema di raccolta e utilizzo dei dati personali possa essere valutato anche alla luce del principio di **equità.** Ad esempio, fermi restando i principi di pertinenza e non eccedenza previsti dalla regolazione, il principio di equità potrebbe essere inteso come grado di condivisione del valore economico che si genera dall'utilizzo dei dati personali. Viene fatto osservare da taluni commentatori, infatti, come tale distribuzione sia verosimilmente "sbilanciata" a vantaggio delle imprese. In altri termini, anche laddove gli utenti ottengano un beneficio netto dall'interazione con il fornitore di un servizio – la cui erogazione presuppone il trasferimento di dati personali – ci si può chiedere se, in una prospettiva di equità, il consumatore debba anche appropriarsi di una parte dei profitti che il fornitore del servizio estrae grazie all'acquisizione dei dati, anche nella forma di un miglioramento dei servizi che riceve.

In questo contesto, peraltro, i *Big Data* tendono a modificare i termini dell'asimmetria informativa tra consumatore e impresa. Prima dell'era digitale, infatti, l'utilizzo dei dati personali tendeva a incidere sulle asimmetrie informative riducendole, mentre attualmente il loro impiego sembra piuttosto orientato a generarle spostandole a carico dei consumatori. Tale asimmetria informativa tende peraltro ad avere una dimensione intertemporale nella misura in cui il rilascio di dati può dare luogo nell'immediato ad un beneficio (ad esempio il miglioramento del servizio) ma in un secondo momento potrebbe avere possibili ricadute negative per il consumatore, che non è in grado di valutare nel momento in cui effettua la decisione di usufruire del servizio e ben potrebbe non essere consapevole delle conseguenze derivanti dalla perdita sul controllo dei propri dati (cfr. *infra*§*5.3.5.*).

5.3.3. La raccolta e l'utilizzo dei dati personali come variabile economica

Poiché non si osserva di norma un mercato nel quale si realizzano autonome relazioni di scambio tra domanda e offerta che abbiano nello specifico ad oggetto i dati personali, viene in rilievo la necessità

[202] Cfr. le modifiche introdotte al Regolamento UE sullo statuto e il finanziamento dei partiti politici europei e delle fondazioni politiche europee dal Regolamento (UE, EURATOM) 2019/493 del Parlamento europeo e del Consiglio che modifica il Regolamento (UE, EURATOM) n. 1141/2014 del Parlamento europeo e del Consiglio, per quanto riguarda la procedura di verifica relativa alle violazioni delle norme in materia di protezione dei dati personali nel contesto delle elezioni del Parlamento europeo (GUUE 27.3.2019). In base alle dette modifiche, se l'Autorità per i partiti politici europei e le fondazioni politiche europee viene a conoscenza di una decisione di un'autorità nazionale di controllo sulla protezione dei dati da cui sia possibile evincere che la violazione delle norme applicabili in materia è connessa ad attività volte ad influenzare deliberatamente o tentare di influenzare l'esito delle elezioni europee, è tenuta ad avviare un'apposita procedura di verifica – anche coordinandosi con l'autorità nazionale di controllo interessata – all'esito della quale possono essere applicate le sanzioni previste nei confronti dei partiti europei o delle fondazioni politiche europee che abbiano utilizzato a proprio vantaggio tale violazione. Le sanzioni potrebbero ammontare al 5% del bilancio annuale del partito o della fondazione interessati. Inoltre, i partiti e le fondazioni che risulteranno aver commesso una violazione non potranno chiedere finanziamenti a carico del bilancio generale dell'Unione europea nell'anno in cui la sanzione è imposta.

di tenere comunque conto del loro ruolo nei mercati in cui gli utenti, nell'acquistare un servizio (primario), acconsentono all'utilizzo dei propri dati da parte delle imprese.

Nell'ottica di far emergere lo "scambio dei dati" che non si realizza in un contesto di mercato "autonomo" e di analizzare gli effetti che esso produce sul benessere sociale e del consumatore, nella teoria economica è possibile considerare la fornitura dei dati personali come una componente implicita del prezzo che l'utente paga all'impresa per l'acquisto del servizio primario o in alternativa una componente qualitativa del servizio.

La possibilità di inquadrare il grado di utilizzo di dati personali come una componente del prezzo o della qualità del servizio presuppone una consapevolezza del verificarsi di un trasferimento dei propri dati personali e del valore economico intrinseco agli stessi, consapevolezza che tuttavia non appare così scontata, come emerge dai risultati della *survey* condotta dall'Autorità (*cfr. infra,*§*5.3.5.*).

Raccolta e utilizzo dei dati personali come prezzo non monetario. Per quanto l'inquadramento di dati come corrispettivo dei servizi possa apparire in conflitto con la filosofia sottostante gli obiettivi di protezione della *privacy*, i dati personali acquisiti dal fornitore di un servizio sono spesso visti dalla letteratura economica e persino dalla più recente giurisprudenza come il prezzo che un utente paga per la fruizione di quest'ultimo[203]. Tale analogia, da un lato, ha il pregio di evidenziare come i dati personali costituiscano di fatto il principale se non l'unico valore di scambio del servizio, laddove quest'ultimo viene erogato dall'impresa gratuitamente.

Dall'altro lato, assimilare la fornitura dei dati ad un prezzo appare riconoscere implicitamente che essa determina una disutilità per l'utente al pari di un esborso monetario e deve in ogni caso tenere conto della natura di diritto fondamentale del diritto alla protezione dei dati personali.

Si tratta, ad esempio, di un approccio che è stato sviluppato anche nell'ambito dell'*enforcement* delle norme a tutela dei consumatori. Infatti, nonostante gli operatori digitali spesso non esigono dall'utente alcun esborso monetario in cambio dei servizi offerti, è possibile configurare, in ogni caso, un rapporto di consumo laddove gli stessi utenti mettono a disposizione della piattaforma e, attraverso questa, di terzi una mole ingente di informazioni collegata al proprio *account*, inclusi i dati personali e quelli dei propri contatti in rubrica. Un siffatto patrimonio informativo, utilizzato, come noto, per la profilazione degli utenti a uso commerciale e per finalità di *marketing*, acquista, in ragione di tale uso, un valore economico che costituisce evidentemente la controprestazione del servizio fornito dalla piattaforma in assenza di corrispettivo monetario. La stessa Commissione europea riconosce che i dati personali, le preferenze dei consumatori e altri contenuti generati dagli utenti hanno un valore economico *de facto* e che una piattaforma che si qualifica come 'professionista' deve sempre rispettare le norme dell'UE in materia di diritto commerciale e dei consumatori nell'ambito delle proprie pratiche commerciali. La commercializzazione di tali prodotti come 'gratuiti' senza informare i consumatori del modo in cui saranno utilizzati i dati relativi alle loro preferenze, i dati personali e i contenuti generati dagli utenti in alcune circostanze può essere considerata, al di là di eventuali ulteriori profili di violazione della disciplina di protezione dei dati, una pratica ingannevole (cfr.

[203] Cfr. European Data Protection Supervisor, Opinion 8/2016 - *EDPS Opinion on coherent enforcement of fundamental rights in the age of big data*, 23 September 2016. Per una rassegna della letteratura relativa ai mercati senza prezzo e al ruolo dei dati come "mezzo di pagamento" si rinvia, tra gli altri, a OCSE (2018), *Quality Considerations in Digital Zero-Price Markets*, *Background note by the Secretariat*, Parigi, 28 novembre. Cfr. altresì la recente pronuncia del Tar Lazio, sentenza 10 gennaio 2020, n. 261, *Facebook*.

provv. AGCM PS11112 - *Facebook-Condivisione dati con terzi*, 29 novembre 2018 n. 27432, su cui v. *infra*,§5.4.1).

Al contempo, come si è rilevato sopra, nei mercati digitali non è sempre univoca la relazione tra fornitura dei dati e benessere dei consumatori. Inoltre, mentre i consumatori hanno di norma piena consapevolezza del prezzo dei beni/servizi che consumano, il livello di *privacy* associato al consumo di determinati beni/servizi costituisce uno degli aspetti probabilmente meno immediatamente percepibili e "quantificabili" dal consumatore.

Protezione dei dati personali come qualità dei servizi[204]**.** In una diversa prospettiva, la protezione dei dati personali può anche essere considerata come una dimensione qualitativa, tra le tante, di un servizio, cosicché, a parità di prezzo ed eventualmente di altre caratteristiche, i consumatori (correttamente informati) dovrebbero tendere a scegliere il servizio che garantisce la minore fornitura possibile di dati o, comunque, un più elevato potere di controllo sui propri dati.

Più in generale, è possibile individuare una varietà di dimensioni connesse al trattamento dei dati personali che possono rilevare, oltre che ai sensi della normativa contenuta nel RGPD, anche in base a una lettura della protezione dei dati come qualità di un servizio. Infatti, si può avere riguardo a: *i)* la tipologia e il volume di dati raccolti; *ii)* la finalità per la quale i dati sono raccolti; *iii)* la durata del trattamento; *iv)* l'eventuale condivisione dei dati con terze parti; *v)* la possibilità per gli utenti di accedere, modificare, cancellare ed esportare i propri dati personali; *vi)* il legame tra i dati raccolti e la possibilità o meno di utilizzare il servizio; *vii)* la trasparenza nella relazione con l'utente in merito alla raccolta e al trattamento di dati personali.

Assimilare il grado di *privacy* alla qualità mette meno in evidenza il ruolo che i dati personali hanno come "valore di scambio" tra consumatore e fornitore del servizio. Tuttavia, associare la *privacy* a una componente qualitativa del servizio può per altri versi evidenziare taluni aspetti delle modalità di valutazione che i consumatori fanno del grado di protezione dei loro dati da parte dell'impresa e del funzionamento dei mercati (cfr. *infra*).

Si tratta, infatti, di un'analogia che appare coerente con un'impostazione che vede il consumatore attribuire alla tutela dei dati personali a sé riferibili anche un valore economico: come per altre caratteristiche qualitative, un livello maggiore di *privacy*, a parità di altre condizioni, dovrebbe corrispondere ad una maggiore utilità per il consumatore.

Tuttavia, occorre considerare che in alcune situazioni e per alcuni servizi digitali il legame tra dati personali e qualità del servizio è più articolato, nella misura in cui la qualità del servizio nel suo complesso dipende dalle informazioni personali che l'impresa è in grado di acquisire sul singolo utente. Spesso, infatti, i servizi che comportano l'estrazione di dati personali riescono a garantire una qualità tanto migliore quanto più ampio è il *set* di dati che l'utente è disposto a trasmettere al fornitore (si pensi, ad esempio, ai servizi di ricerca e di *matching*, in cui la personalizzazione costituisce un elemento fondamentale della qualità).

Va in ogni caso ricordato che, attesa la natura di diritto fondamentale della protezione dei dati personali, gli ambiti di "negoziabilità" da parte dell'interessato devono trovare modalità di

[204] Cfr. OCSE (2018), *Quality Considerations in Digital Zero-Price Markets*, *Background note by the Secretariat*, Parigi, 28 novembre, e la copiosa letteratura ivi citata; e OCSE, 2018, *Considering non-price effects in merger control*, DAF/COMP(2018)2. Si veda altresì la Decisione della Commissione europea, C(2016) 8404, Case M.8124 – Microsoft/LinkedIn, 6 Dicembre 2016, http://ec.europa.eu/competition/mergers/cases/decisions/m8124_1349_5.pdf.

espressione nelle forme consentite dalla relativa disciplina, anzitutto rispettando i principi fondanti di protezione dei dati personali, compendiati all'art. 5 del RGPD, oltre che riscontrando la sussistenza di una delle condizioni di liceità del trattamento indicate al successivo art. 6.

5.3.4. La relazione tra concorrenza e utilizzo dei dati personali

Allo scopo di esaminare la relazione tra concorrenza e utilizzo dei dati personali, analizzando qual è l'impatto della pressione concorrenziale tra i fornitori di un servizio sulla quantità di dati personali che essi estraggono presso gli utenti, occorre innanzitutto chiarire come la domanda e l'offerta di dati personali influenzano la domanda e l'offerta del servizio primario. Impregiudicata la natura di diritto fondamentale della protezione dei dati personali e l'impossibilità di qualificare tali dati come "merce"[205], la "fornitura dei dati", che non emerge esplicitamente in un mercato "autonomo", costituisce un elemento essenziale sul quale di fatto si basano i mercati digitali dei servizi primari.

Considerare il grado di utilizzo dei dati personali come un aspetto del prezzo che i consumatori pagano per un bene/servizio ovvero come una dimensione qualitativa di quest'ultimo costituisce elemento utile per analizzare il legame tra concorrenza e sfruttamento dei dati personali, pur nella consapevolezza della natura di diritto fondamentale del diritto alla protezione dei dati personali e delle regole poste a suo presidio[206].

L'assimilazione dell'utilizzo dei dati personali a un prezzo implica, come è già stato chiarito in precedenza, che gli utenti abbiano la capacità di attribuire consapevolmente alla fornitura dei propri dati un valore economico puntuale. Sotto una diversa prospettiva, tale analogia sembra suggerire l'esistenza di un chiaro rapporto tra concorrenza e utilizzo dei dati personali: la maggiore richiesta di dati, così come l'incremento del prezzo, infatti dovrebbe essere minore nei mercati nei quali è maggiore la pressione concorrenziale esercitata dalle imprese. In quest'ottica, un contesto più favorevole alla concorrenza potrebbe altresì ridurre i rischi di discriminazione e le implicazioni in termini di eguaglianza sostanziale, frutto della diversa disponibilità a sostituire la fornitura dei dati personali con un pagamento monetario.

Tuttavia, quando la tutela della *privacy* viene assimilata alla qualità, emerge più chiaramente la complessa relazione tra pressione concorrenziale nel mercato dei servizi primari e grado di utilizzo dei dati personali.

In questo contesto, si può ragionevolmente ipotizzare che gli utenti abbiano una maggiore o minore preferenza per la fornitura dei propri dati personali e le imprese differenzino i servizi che offrono anche sulla base del grado di acquisizione e utilizzo di dati personali. Peraltro, come detto, una minore richiesta di dati da parte della piattaforma non sempre si associa ad un livello qualitativo complessivo del servizio superiore, dal momento che in talune circostanze l'offerta di un servizio di qualità elevata richiede la raccolta e l'elaborazione di una considerevole mole di dati personali, di talché la relazione tra grado di utilizzo dei dati personali e qualità generale del servizio diventa ambigua.

La scelta del consumatore è dettata da diversi fattori che dipendono dal grado desiderato di utilizzo di dati personali da parte del fornitore del servizio, dalla qualità generale del servizio e dal suo prezzo. A seconda di come si compone il *trade-off* tra queste variabili si possono osservare scelte diversificate

[205] Cfr. Direttiva (UE) 2019/770 del 20 maggio 2019, *relativa a determinati aspetti dei contratti di fornitura di contenuto digitale e di servizi digitali* (v. *supra*, nota 121).
[206] Cfr. l'assunto, sopra richiamato, della Corte di Giustizia (Grande Sezione), sentenza del 13 maggio 2014 nella causa C-131/12, *Google Spain*, par. 97.

degli utenti, situazione che a sua volta dovrebbe condurre le imprese a offrire servizi che presentano livelli di qualità, e dunque livelli di utilizzo di dati personali, differenziati.

Nei mercati digitali, tuttavia, non sembra rilevarsi frequentemente un'elevata **differenziazione del grado di utilizzo dei dati personali** all'interno delle offerte delle singole imprese. In altri termini, i grandi operatori digitali non offrono di norma ai consumatori un insieme di opzioni con diversi livelli di utilizzo dei dati personali (a cui possono corrispondere diverse funzionalità e/o prezzi monetari). Pertanto, la scelta che l'utente compie si traduce nella selezione del fornitore, laddove gli operatori tendono ad offrire servizi relativamente differenziati, o eventualmente nella scelta di non fruire del servizio, se gli operatori offrono servizi tendenzialmente omogenei e il mercato non è maturo. Esistono tuttavia alcune eccezioni all'assenza di differenziazione nell'offerta di servizi da parte del singolo operatore: in alcuni casi, infatti, i portafogli di servizi offerti da taluni operatori garantiscono un grado più elevato di tutela dei dati personali agli utenti che acquistano versioni "premium", generalmente a pagamento, del servizio in questione.

Ad ogni modo, anche la differenziazione tra operatori non appare particolarmente elevata: infatti, i fornitori di servizi digitali tendono ad offrire servizi con caratteristiche simili, in termini di prezzo e grado di utilizzo dei dati personali, proponendo in particolare servizi il cui utilizzo è spesso associato a un'ampia raccolta e utilizzo di dati personali. Non mancano comunque alcune eccezioni, anche di particolare rilievo, ossia operatori che hanno reso la (maggiore) tutela dei dati personali degli utenti un elemento qualificante della propria offerta commerciale: in questa direzione va l'offerta commerciale di Apple e lo sviluppo di servizi di ricerca e di *browser*, come DuckDuckGo, che si dichiarano particolarmente attenti alla *privacy* degli utenti.

In generale, riconoscere che il grado di utilizzo dei dati personali può costituire un aspetto di differenziazione delle offerte commerciali disponibili per gli utenti significa anche riconoscere che i consumatori potrebbero godere di livelli rafforzati di *privacy* (e di un minor quantitativo di dati forniti) solo se hanno la disponibilità a pagare – ad esempio in termini monetari – per i servizi che fanno un utilizzo più limitato dei dati personali degli utenti. In tale scenario, alcuni rilevano il rischio che la fruizione di servizi digitali che non comporti un elevato grado di raccolta e utilizzo dei dati personali degli utenti possa diventare "un lusso per pochi".

Al di là delle possibili differenziazioni sul grado di utilizzo dei dati personali tra le diverse versioni di un servizio offerte da un'impresa e tra imprese diverse, dunque, una questione rilevante è se la concorrenza costituisce strumento idoneo a garantire agli utenti un livello minimo di *privacy*[207].

In linea generale, la risposta a tale interrogativo appare connessa a due aspetti principali: *i)* il grado di eterogeneità tra le preferenze degli utenti in merito alla disponibilità a fornire i propri dati personali; *ii)* i costi "impliciti" che le imprese potrebbero dover sostenere per offrire una maggiore *privacy* (intesa anche come minore richiesta di dati) agli utenti in termini, ad esempio, di una riduzione della qualità complessiva del servizio offerto (agli utenti stessi o ad operatori attivi su un altro versante della piattaforma).

Ad esempio, è ragionevole attendersi che anche servizi associati con un elevato grado di acquisizione e utilizzo di dati personali degli utenti possono sopravvivere in un mercato concorrenziale laddove vi siano consumatori che non attribuiscono un valore elevato alla condivisione dei propri dati personali

[207] Cfr. Michael L. Katz, 2019, Multisided Platforms, Big Data, and a Little Antitrust Policy, *Review of Industrial Organization,* 54:695–716.

con le imprese. Al contempo, appare ragionevole attendersi che il mercato possa sostenere un grado elevato di raccolta e utilizzo di dati personali (pur sempre nel rispetto della relativa cornice normativa) laddove i dati personali siano particolarmente rilevanti per offrire un bene/servizio di qualità complessiva più elevata.

Per altro verso, l'emergere di servizi con modelli di *business* volti a limitare la raccolta e l'utilizzo di dati personali in un mercato concorrenziale richiede che vi siano consumatori disposti a "pagare" anche per servizi con tali caratteristiche, in termini monetari ovvero implicitamente (ad esempio, consumando servizi di qualità complessiva peggiore)[208].

5.3.5. Domanda e offerta di dati personali

In generale, al fine di disporre di un approccio analitico utile all'analisi della *performance* dei mercati, può essere utile considerare domanda e offerta di dati personali "direttamente" e non solo come un aspetto della concorrenza nel mercato del bene/servizio primario[209].

La **domanda di dati personali** è espressa dalle imprese, le quali hanno sempre un incentivo ad acquisirli, nella misura in cui la loro raccolta ed elaborazione ne accresce i profitti, che possono derivare da diverse fonti a seconda del modello di *business* adottato.

Ad esempio, maggiori profitti possono essere realizzati dalle imprese non soltanto nell'ambito della fornitura del servizio primario, nella misura in cui i dati personali possono servire a migliorare, attraverso una personalizzazione, l'attrattività del servizio rivolto ai consumatori, ma anche nell'offerta all'utente di servizi aggiuntivi a pagamento o per realizzare ricavi dalla cessione dei dati a terzi.

Sussiste dunque un incentivo ad acquisire quanti più dati possibili da parte delle imprese, in quanto, a fronte di incremento dei ricavi che possono realizzare aumentando gli utilizzi dei dati raccolti, esse hanno la possibilità di sfruttare le economie di scala e di scopo che caratterizzano la loro estrazione ed elaborazione. Da un lato, infatti, i costi derivanti dall'elaborazione dei dati, essendo perlopiù riconducibili ad investimenti in infrastrutture di calcolo, sono prevalentemente di natura fissa (quantomeno all'interno di intervalli prefissati di capacità di calcolo). Dall'altro lato, l'incidenza dei costi di raccolta e di elaborazione dei dati diminuisce all'aumentare del loro utilizzo da parte delle imprese.

Inoltre, laddove la raccolta dei dati personali sostiene un modello di *business* configurabile come una piattaforma a due versanti, gli effetti di rete indiretti, e dunque le esternalità positive che si instaurano tra i due versanti del mercato, costituiscono altresì un ulteriore impulso all'estrazione dei dati da parte delle imprese. In particolare, tale modello di *business* determina sia l'incentivo ad estendere la raccolta dei dati su una platea più ampia di utenti, in quanto ciò favorisce un ampliamento della domanda sull'altro versante della piattaforma, sia l'incentivo ad intensificare l'attività di estrazione

[208] Cfr. Norman, G. et al (2016), "Competition and consumer data: The good, the bad, and the ugly", 70 *Research in Economics* 4, https://www.sciencedirect.com/science/article/pii/S1090944316301752.
[209] Cfr. Acquisti, A. (2010), "The Economics of Personal Data and the Economics of Privacy", *Background Paper for OECD Working Party for Information Security and Privacy and Working Party on the Information Economy*, www.oecd.org/sti/ieconomy/46968784.pdf; Farell, J. (2012), "Can privacy be just another good?", *Journal on Telecommunications and High Technology Law*, Vol. 10, www.jthtl.org/content/articles/V10I2/JTHTLv10i2_Farrell.PDF.

dei dati sulla propria base utenti, in quanto ciò potrebbe aumentare il valore economico del *set* di dati personali relativo ai singoli utenti che cede all'altro versante del mercato.

L'**offerta di dati personali** è per contro espressa dagli utenti, per i quali, come detto, la relazione tra fornitura dei dati e utilità non è univoca. La fornitura dei dati genera infatti dei costi in capo all'utente, che però possono essere almeno in parte controbilanciati dalla maggiore qualità o dal minor prezzo a cui il servizio primario viene offerto.

La circostanza che le preferenze degli utenti sul grado di tutela dei propri dati personali emergano nella scelta relativa alla fruizione di servizi primari comporta che condizione necessaria (ma non sufficiente) perché la concorrenza generi un livello ottimale di utilizzo di dati personali è che gli utenti, nell'esprimere la propria domanda di servizi primari tengano adeguatamente in considerazione il diverso livello di *privacy* che essi assicurano.

Nell'assetto attuale dei mercati digitali sono invece riscontrabili delle rilevanti asimmetrie informative tra venditori e consumatori e distorsioni comportamentali che compromettono la capacità degli utenti di scegliere i servizi tenendo conto del loro livello di tutela della *privacy*[210]. Pertanto, l'assenza di consapevolezza dell'utente impedisce che la *privacy* possa costituire una dimensione del processo di interazione tra le imprese, suscettibile di essere disciplinata dalla concorrenza nel mercato dei servizi primari.

Sul punto, appare utile osservare che anche il campione di utenti preso in esame nell'ambito della *survey* condotta dall'AGCM non mostra una piena consapevolezza dell'attività di estrazione dei propri dati personali sottostante la fruizione dei servizi *on line*. Infatti, come sopra illustrato (§ 5.1) circa il 40% degli utenti intervistati ha mostrato di non avere consapevolezza né del fatto che la navigazione in internet e l'utilizzo di app e servizi *online* comporti la raccolta di dati personali né del fatto che tali dati possano essere ceduti dal fornitore del servizio a terzi[211].

Inoltre, si è osservato che anche le scelte di utenti consapevoli possono soffrire di distorsioni altrettanto critiche. Si tratta innanzitutto del c.d. "*privacy paradox*"[212], situazione che si caratterizza per il fatto che, seppure i consumatori esprimano un grande interesse per la tutela della *privacy* e la considerino un importante fattore della qualità di un servizio, gli stessi non sembrano, tuttavia, effettuare scelte di consumo coerenti con tale preferenza dichiarata. Tale situazione si riscontra ad esempio quando un utente non rinuncia a fruire di un servizio gratuito anche laddove quest'ultimo garantisce un livello di tutela della *privacy* molto basso e altri servizi presenti sul mercato offrono, a fronte di un prezzo positivo, una maggiore protezione dei dati personali. Si genera dunque un divario tra preferenze dichiarate e preferenze rivelate dall'utente che può determinare una tutela della *privacy* sub-ottimale.

L'esistenza di questo effetto sembra potersi inferire anche dai risultati della *survey*, ed in particolare dalle risposte alle domande 11 e 6.2. Mentre le risposte alla prima mostrano che il 93% circa degli

[210] Cfr. Acquisti, A., K. Brandimarte, G. Loewenstein (2015), "Privacy and human behaviour in the age of information", *Science*, Vol. 347, http://science.sciencemag.org/content/347/6221/509.
[211] Cfr. Risposte alle domande 1.1 e 1.2.
[212] Cfr. Patricia A. Norberg, Daniel R. Horne, David A. Horne, 2007, The Privacy Paradox: Personal Information Disclosure Intentions versus Behaviors, *The Journal of Consumer Affairs*, 41(1): 100-126.

intervistati ha interesse a tutelare la propria *privacy*[213], la seconda evidenzia che di fatto solo un terzo rifiuta il consenso all'acquisizione e all'utilizzo dei dati.

Un'altra distorsione che si rileva nelle scelte di fruizione di servizi da parte di un utente, ed è in parte collegato alla precedente, consiste nel c.d. "*free effect*". Essa attiene al fatto che gli utenti tendono in buona sostanza ad attribuire un valore sproporzionato ai servizi offerti gratuitamente nonostante questi ultimi comportino un deterioramento significativo della qualità di un servizio. In altri termini, la valutazione dei servizi gratuiti da parte dell'utente tende ad essere sproporzionata rispetto al rapporto costi-benefici che essi determinano. Ciò si traduce nel fatto che l'impatto di una riduzione del prezzo sull'utilità del consumatore non è costante, in quanto cresce in corrispondenza del momento in cui un servizio passa dall'avere un prezzo positivo all'essere offerto gratuitamente.

Tale effetto presenta una serie di risvolti sul piano dell'*outcome* che la concorrenza può determinare sul livello di *privacy*. Infatti, anche un piccolo aumento di prezzo che migliora in maniera esponenziale la qualità del servizio da parte di un potenziale entrante potrebbe non essere sufficiente per quest'ultimo per consentirgli l'ingresso sul mercato. Ciò, inoltre, fa emergere come, anche nel caso in cui i servizi si differenzino sulla base del diverso livello di tutela della *privacy* che garantiscono, di fatto il prezzo rimane il principale *driver* nelle scelte di consumo dell'utente. Tendono dunque a configurarsi mercati nei quali la domanda è espressa principalmente o pressoché esclusivamente da utenti che preferiscono servizi offerti gratuitamente a fronte di una tutela della *privacy* molto bassa (rispetto ai costi generati dal rilascio dei dati).

In questo senso, pur non avendo verificato la sussistenza di una simile distorsione nei comportamenti, la *survey* ha messo in evidenza il grande apprezzamento per i servizi gratuiti tra gli utenti intervistati. Infatti, meno di un quarto di essi (circa il 23%) sarebbe disposto a rinunciare alla fruizione di servizi gratuiti e una percentuale ancora inferiore (circa il 10%) sarebbe disposta a pagare per evitare la raccolta dei dati personali.

In generale, dunque, le distorsioni che si osservano nei comportamenti dei consumatori generano delle vischiosità sul mercato che non consentono agli utenti di massimizzare l'utilità che essi estraggono dalla fruizione dei servizi, comportando il rischio che il livello di tutela della *privacy* generato in equilibrio sul mercato dall'interazione tra domanda e offerta non sia ottimale dal punto di vista del benessere sociale, ed in particolare dal punto di vista di quello del consumatore.

5.3.6. *Privacy*, funzionamento dei mercati e il ruolo della politica pubblica

Le considerazioni di cui sopra suggeriscono l'esistenza di due principali ostacoli al funzionamento di un meccanismo di mercato idoneo a generare un livello di utilizzo dei dati personali "soddisfacente" per gli individui e la società: *i)* la limitata percezione e consapevolezza da parte dei consumatori in merito alla raccolta e all'utilizzo dei propri dati personali da parte dei fornitori di servizi; *ii)* il limitato grado di pressione concorrenziale che caratterizza taluni servizi digitali.

In questo scenario, sono diversi gli strumenti di politica pubblica che in linea di principio possono essere utilizzati per perseguire, direttamente e/o indirettamente, l'obiettivo di un'adeguata tutela della *privacy* nell'ecosistema digitale.

[213] Tale interesse si inferisce dal fatto che gli utenti, posti di fronte all'alternativa se acquistare un servizio base – cui corrisponde la raccolta di un insieme limitato dei dati - o la versione premium –consentendo una maggiore estrazione di dati- per oltre il 90% non acquisterebbero il servizio premium.

Diritti di proprietà e architetture alternative di gestione dei dati. Una prima questione inerisce alla definizione di possibili diritti di proprietà sui dati[214]. Ad avviso di alcuni economisti, infatti, il riconoscimento in capo all'individuo di un diritto di proprietà sui propri dati costituirebbe il presupposto necessario per consentire l'instaurarsi di dinamiche di mercato[215].

Una seconda questione è connessa alle architetture centralizzate attualmente utilizzate per la gestione dei dati. Nel corso dell'Indagine conoscitiva, da alcuni intervenienti è emerso che il modello di piattaforma *online* affermatosi è caratterizzato da una gestione centralizzata dei dati da parte dei titolari delle piattaforme mentre gli utenti, cui i dati si riferiscono, non hanno cognizione di come e dove i dati siano registrati e conoscono le modalità e le finalità del trattamento solo nei limiti di quanto reso noto dai medesimi gestori delle piattaforme. Gli utenti sono estranei al trattamento dei loro dati ed esercitano i diritti che sono loro riconosciuti dalla normativa sulla protezione dei dati personali in maniera mediata, attraverso la piattaforma.

Un modello alternativo potrebbe essere quello in cui ciascun utente controlla i dati che ad esso si riferiscono e decide in autonomia se e in che misura condividerli con le piattaforme che di volta in volta avanzano richiesta in tal senso ("modello decentrato").

Tuttavia, al di là dell'esistenza di diritti di proprietà sui dati e/o di piattaforme decentralizzate per la loro gestione, lo sviluppo di un mercato per lo scambio di dati personali incontra comunque dei limiti che sono principalmente riconducibili al fatto che l'utente generalmente non è in grado di attribuire un valore economico ai suoi dati personali, e quindi non è in grado di individuarne il relativo "prezzo di cessione". Ciò viene in particolare rilievo anche tenuto conto della circostanza che il valore dei dati si forma in una sequenza di passaggi di proprietà e di utilizzi che *ex ante* l'utente non è in grado di prevedere. In questo contesto, appare difficile poter identificare una precisa relazione tra livello di fornitura dei dati ed entità dell'esborso monetario. L'unica valutazione che il consumatore è in grado di fare riguarda l'eventuale presenza di un prezzo implicito – derivante dall'attività di raccolta dei dati da parte del fornitore del servizio primario – e al più un confronto relativo tra i prezzi impliciti che i servizi erogati dai diversi fornitori comportano. Quest'ultima situazione, tuttavia, presuppone che i termini contrattuali dei diversi servizi siano effettivamente confrontabili.

Disciplina di protezione dei dati personali, portabilità e circolazione dei dati. Si è detto che l'acquisizione dei dati è rilevante per le imprese nella misura in cui consente loro di estrarre informazioni che permettono di realizzare un bene/servizio maggiormente competitivo. In questa fase, il rispetto della disciplina in materia di protezione dei dati personali assume un ruolo fondamentale e se, da un lato, rappresenta il presupposto per tutelare adeguatamente i dati personali dell'interessato, dall'altro può rendere più difficoltoso l'accesso ai dati da parte degli operatori che non beneficiano di un rapporto diretto con l'utente. La protezione dei dati da parte dei relativi titolari si scontra infatti con l'esigenza di incentivare la circolazione dei dati stessi e, con essa, la libera concorrenza tra le imprese.

[214] Cfr. audizione del Prof. Gambaro (18 dicembre 2017).

[215] Cfr. audizione dei Proff. Preta, Maggiolino e altri (30 gennaio 2018). Cfr. altresì Robert Bartlett, Developments in the Law—The Law of Cyberspace, 112 *Harv. L. Rev*. 1574 (1999); Lawrence Lessig, The Architecture of Privacy, 1 *Vand. J. Ent. L. & Prac.* 56 (1999); Richard S. Murphy, Property Rights in Personal Information: An Economic Defense of Privacy, 84 *Georgetown L.J.* 2381 (1996).

La normativa rilevante prevede sia un quadro generale, individuato nel RGPD, sia regole speciali relative al trattamento dei dati personali e alla tutela della vita privata nel settore delle comunicazioni elettroniche (attualmente in corso di aggiornamento).

Pure definendo la cornice entro la quale i dati personali possono essere utilizzati dal titolare del trattamento, la regolamentazione lascia ampi spazi di autonomia per queste ultime, che possono rilevare sia per l'applicazione della normativa a tutela del consumatore che della normativa *antitrust*.

Tra i diritti più significativi previsti dal RGPD rileva il menzionato diritto alla portabilità dei dati (art. 20), il quale assolve a due scopi: aumentare il controllo dell'interessato sui suoi dati personali e facilitare la trasmissione dei dati da un operatore ad un altro. Esso appare senza dubbio un punto di raccordo tra la disciplina di tutela dei dati personali e quella in materia di concorrenza, trattandosi di una situazione giuridica soggettiva – il cui esercizio rimane saldamente nelle mani dell'interessato e la cui violazione può formare oggetto di sindacato avanti all'autorità di protezione dei dati[216] - idonea a produrre anche significativi effetti pro-concorrenziali, in termini sia di circolazione dei dati che di mobilità degli utenti.

A sua volta il diritto alla portabilità dovrebbe contribuire a scongiurare il *lock-in* tecnologico e ad aumentare la concorrenza tra le imprese che forniscono servizi digitali[217]. Ad esempio, per attrarre consumatori che si servono di un determinato operatore di servizi digitali un concorrente potrebbe offrire servizi di integrazione dei loro dati personali (o di parti di essi) (si pensi ai *social network* che incorporano i contatti degli utenti tramite l'API di contatto del *provider* di posta elettronica[218]).

Sussistono tuttavia diversi ostacoli all'effettivo sviluppo della portabilità, legati in particolare alla scarsa consapevolezza degli utenti circa l'esistenza di tale diritto, ai vincoli alla loro mobilità (dovuti anche alla presenza di esternalità di rete) e ai confini ancora incerti della portabilità, che include soltanto una parte dei dati a disposizione del titolare del trattamento.

La *survey* condotta ha rilevato che gli utenti, a qualche mese dell'entrata in vigore del RGPD, per la gran parte (91%) non erano a conoscenza dei diritti all'accesso e alla portabilità dei dati. A questo riguardo, affinché la portabilità sia concretamente applicata, è necessario che gli utenti siano adeguatamente informati circa il valore dei propri dati e il loro potenziale utilizzo e, allo stesso tempo, che non "si impigriscano" di fronte alla possibilità di un cambio troppo complesso[219].

Quanto ai vincoli allo spostamento, gli utenti possono risultare poco propensi a esercitare il diritto alla portabilità e a spostarsi da una piattaforma all'altra in assenza di valide alternative[220]: incentivare

[216] In tale prospettiva, cfr. le Linee Guida sul diritto alla portabilità dei dati – WP242 adottate dal Gruppo di lavoro Art. 29 il 13 dicembre 2016 e poi emendate, a seguito di consultazione pubblica, e adottate il 5 aprile 2017.

[217] Da questo punto di vista l'art. 16, par. 4, secondo periodo, della Direttiva (UE) n. 2019/770 del 22 maggio 2019 relativa a determinati aspetti dei contratti di fornitura di contenuto digitale e di servizi digitali, si distingue dall'art. 20 RGPD. Nell'ambito della Direttiva, il diritto alla portabilità sorge solo in relazione a contenuti diversi dai dati personali, che sono stati forniti o creati dal consumatore utilizzando il contenuto digitale o il servizio digitale, e solo dopo che il consumatore ha risolto il contratto. La Direttiva, dunque, tutela piuttosto il diritto di recesso del consumatore al fine di evitare effetti di blocco. In particolare, le norme servono ad evitare che il timore dell'utente che il contenuto possa essere perso con il recesso dal contratto abbia un ruolo nella decisione di esercitare tale diritto. Parimenti, anche il Regolamento UE 2017/1128 sulla portabilità *cross border* dei contenuti *online* ha un altro *focus,* in quanto le norme sulla portabilità transfrontaliera mirano a garantire che i contenuti digitali acquisiti da un consumatore in uno Stato membro siano accessibili gratuitamente da qualsiasi altro Stato membro.

[218] V., tuttavia, *infra* nel testo esempi di come la disciplina del diritto alla portabilità dei dati non faciliti questo effetto pro-concorrenziale.

[219] Cfr. audizione dei Proff. Preta, Maggiolino e altri (30 gennaio 2018).

[220] Cfr. audizione del Prof. Mantelero.

gli utenti ad esercitare tale diritto potrebbe rivelarsi particolarmente oneroso per le imprese nuove entranti, che non possono contare su di una base di utenti vasta come quella delle imprese più strutturate e su di un rapporto diretto con gli utenti.

Va infine considerato che l'ambito di applicazione del diritto alla portabilità risulta comunque limitato al dato c.d. *fornito* e non alle informazioni *estratte* dal dato. Le Linee guida sul diritto alla portabilità dei dati, adottate dal gruppo di lavoro art. 29 per la protezione dei dati (WP29), hanno chiarito che la nozione di dati "forniti" da un interessato debba riferirsi ai dati forniti consapevolmente e attivamente dall'interessato, ai dati personali osservati sulla base delle attività svolte dagli utenti, come per esempio i dati grezzi generati da un contatore intelligente, la cronologia della navigazione su un sito web o la cronologia delle ricerche effettuate. Restano tuttavia esclusi i dati "generati" dal titolare (utilizzando come input i dati osservati o forniti direttamente), come ad esempio il profilo-utente creato a partire dall'analisi dei dati grezzi[221].

Di conseguenza, i dati inferenziali e derivati creati dal titolare sulla base dei dati forniti dall'interessato, anche se a lui riferiti o riferibili, quali i risultati prodotti da un algoritmo, esulano dal campo di applicazione del diritto alla portabilità, in quanto spesso frutto di analisi basate su tecniche di *data analysis* che rientrano nel patrimonio informativo del titolare e che possono essere coperte da diritti di proprietà intellettuale.

Questo limite, pur apparentemente chiaro e condivisibile, può in concreto dare luogo a difficoltà applicative, laddove non risulti agevole distinguere tra dati forniti anche indirettamente o involontariamente dall'interessato e dati che sono frutto di un'autonoma elaborazione del loro titolare.

Ulteriore limite specifico all'esercizio del diritto alla portabilità è la sua fattibilità tecnica. L'articolo 20, par. 2, del RGPD obbliga infatti il titolare a trasmettere i dati portabili direttamente a un diverso titolare "*se tecnicamente fattibile*". Secondo le linee guida WP29, la fattibilità tecnica della trasmissione da un titolare all'altro dovrà essere valutata caso per caso, tenendo tuttavia presente che ciò che la norma richiede è che i sistemi siano *interoperabili* e non necessariamente *compatibili* (cfr. considerando 68 del RGPD). La previsione potrebbe dar luogo a facili strumentalizzazioni e occorre evitare che un rifiuto di adempiere alla richiesta di portabilità sia motivato dalle caratteristiche di un formato non interoperabile o dalle modalità di trattamento adottate del titolare e non piuttosto da effettivi impedimenti tecnici o inadeguatezza del sistema ricevente.

Per evitare che gli utenti possano avvalersi del diritto alla portabilità solo in limitate circostanze, potrebbe essere centrale lo sviluppo di *standard* comuni di trasferimento dati[222].

Nell'ambito dell'Indagine alcuni operatori hanno evidenziato come, al fine di favorire la libera circolazione dei dati e lo sviluppo dell'economia digitale, si potrebbe sviluppare un modello con sistemi decentralizzati e alternativi, nei quali gli utenti abbiano il controllo dei dati generati nelle loro diverse attività e possano decidere in maniera autonoma se e come metterli in comune per uno scopo al quale attribuiscono valore[223]. Lo sviluppo di tali sistemi richiede d'altro canto la definizione di

[221] Cfr. audizione dei Proff. Preta, Maggiolino e altri (30 gennaio 2018).
[222] Inge Graef, "*Data portability at the crossroads of data protection and competition policy*", intervento al convegno "Big data e concorrenza" che si è svolto presso la LUISS Guido Carli il 9 novembre 2016.
[223] Cfr. audizione dei Proff. Giannotti e Pedreschi (5 dicembre 2017) e audizione del dott. Quintarelli (13 settembre 2018).

adeguati strumenti per l'archiviazione personalizzata dei dati, per l'interazione tra utenti e tra utenti e piattaforme e per la trasmissione dei dati in modalità sicura[224].

In proposito, sul piano tecnico le Linee Guida individuano due soluzioni distinte: la portabilità dei dati potrebbe, infatti, avvenire attraverso trasmissione diretta dell'intero insieme di dati portabili (o di parte di essi) all'utente, ad esempio tramite *download*, ovvero consentendo a soggetti terzi di trasmettere i dati direttamente ad altri titolari del trattamento per il tramite di uno strumento automatizzato.

A questo riguardo si rileva che il mercato già offre una varietà di strumenti. Ad esempio, nel 2011 è stato sviluppato *Takeout*, che permette agli utenti di Google di eseguire il *download* di tutti i propri dati forniti per usufruire dei servizi di Google (*email,* foto, calendario, contatti, documenti archiviati in Google Drive, etc.) e di effettuare un *upload* di tali dati su un'altra piattaforma di proprio gradimento. Questa modalità di portabilità, tuttavia, soffre proprio delle limitazioni più sopra descritte: il suo effettivo funzionamento, infatti, è legato alla proattività dell'utente e alle possibilità di archiviazione (la mole di dati scaricati potrebbe essere, infatti, significativa), nonché alla interoperabilità tra piattaforme (si pensi, ad esempio, ai richiamati possibili conflitti tra diversi formati di *file*).

Per questo motivo, ponendosi maggiormente nel solco della seconda soluzione descritta *supra,* il mercato ha iniziato ad offrire sistemi che consentono agli utenti di trasferire i loro dati direttamente da un servizio all'altro anche senza effettuare il *download* e il successivo *upload* dei propri dati. Ad esempio, Google, in partenariato con Microsoft, Twitter e Facebook, nel 2018 ha lanciato il c.d. *Data Transfer Project*, che consiste in un sistema *open source* per la promozione dell'*universal data portability* per permettere agli utenti di esportare dati e importarli su quanti più servizi possibili.

Altri esempi di strumenti di portabilità dei dati attraverso un mezzo automatizzato (generalmente *app*) vengono, *inter alia,* dall'Italia, dove sono stati sviluppati sistemi che consentono agli utenti di richiedere a diversi titolari di trattamento l'estrazione dei propri dati personali e di archiviarli in un'area dedicata (ad esempio, *Weople*). Con l'inserimento dei dati personali nell'area dedicata l'utente riceve una remunerazione, che deriva dai pagamenti che i titolari di trattamento effettuano per avere degli spazi pubblicitari all'interno di dette aree e per la mera lettura delle offerte commerciali da parte degli utenti destinatari (c.d. mercato dell'attenzione). I dati presenti nelle aree dedicate possono essere, peraltro, ulteriormente elaborati tramite un'attività di profilazione. Successivamente i risultati di tali attività verrebbero venduti alle imprese interessate alla loro acquisizione, anche in questo caso per lo svolgimento campagne pubblicitarie e/o offerte commerciali personalizzate, generando per gli utenti ulteriori fonti di remunerazione dell'utilizzo dei dati personali degli utenti medesimi.

Tali iniziative potrebbero fungere da strumento di *consumer empowerment* potenzialmente in grado di superare in parte le descritte limitazioni derivanti dall'attuale disciplina del diritto alla portabilità dei dati, in termini di contribuzione alla costruzione di una consapevolezza da parte degli utenti circa il valore economico dei loro dati personali, grazie all'ottenimento di una remunerazione per l'utilizzo di tali dati da parte di soggetti terzi.

[224] Cfr. audizione dei Proff. Giannotti e Pedreschi (5 dicembre 2017) e audizione del dott. Quintarelli (13 settembre 2018).

Una volta chiariti i confini del diritto alla portabilità, se pur con le menzionate incertezze che solo l'esercizio effettivo del diritto potrà aiutare a superare, residuano le ipotesi in cui il titolare del trattamento, che si trovi in posizione di dominanza, sia tenuto a fornire l'accesso ai dati nella sua disponibilità, al di là di quanto previsto dal diritto alla portabilità. Si tratta di casi eccezionali, dove l'eventuale rifiuto del soggetto dominante a fornire l'accesso ai propri dati può assumere un rilievo dal punto di vista del diritto della concorrenza soltanto se i dati in questione costituiscano elementi essenziali per lo svolgimento dell'attività dell'impresa che richiede l'accesso e siano effettivamente unici e non duplicabili. In tali fattispecie, la normativa in materia di tutela dei dati personali può per certi aspetti confliggere con quella a tutela della concorrenza nella misura in cui, per poter consentire l'accesso ai propri dati, il titolare dovrebbe comunque acquisire il consenso dei soggetti interessati.

Dall'indagine condotta è emerso tuttavia come non sia infrequente che gli utenti neghino il consenso al trattamento dei dati: circa un terzo dei rispondenti (33,4%) ha infatti affermato di aver "spesso" negato il consenso. Tale circostanza rappresenta un ulteriore ostacolo alla circolazione dei dati e, conseguentemente, allo sviluppo di servizi in concorrenza.

5.4. Condotte *data-driven* tra la tutela della concorrenza e la tutela del consumatore

Nel settore dei *Big Data*, condotte potenzialmente rilevanti ai fini dell'applicazione della normativa a tutela della concorrenza e del consumatore possono emergere in tutte le fasi della "filiera", che comprendono sia la raccolta dei dati (cfr. 1.3.1.) che la gestione e l'elaborazione degli stessi per l'offerta di servizi e la loro personalizzazione (cfr. 1.3.2.).

La nozione di ecosistema è utile a cogliere la complessità delle interazioni, dei flussi e degli scambi che si realizzano nel settore digitale, sia nel suo complesso sia nell'ambito dei sistemi di servizi collegati alle principali piattaforme. Tuttavia, anche nel settore digitale, la disciplina analitica che la definizione dei mercati rilevanti impone ha un valore fondamentale, potendo poi il singolo mercato rilevante ben essere considerato anche all'interno del più complesso ecosistema nel quale si colloca. Si tratta di un esercizio complesso, ma che può svilupparsi anche sulla base delle metodologie e degli approcci già consolidati nell'analisi *antitrust* dei mercati a più versanti e degli *aftermarkets*.

Ciò posto, mercati rilevanti nei quali sono state già accertate posizioni dominanti comprendono i servizi di ricerca *online*, i *social network*, nonché i sistemi operativi per dispositivi mobili. Anche a livelli più "profondi" dell'ecosistema, tuttavia, alcuni operatori appaiono detenere posizioni di mercato di particolare rilievo nella fornitura di servizi alle imprese connessi alla raccolta e/o all'elaborazione dei *Big Data* (*analytics*, *cloud computing*, *data storage*).

In tale contesto, vengono affrontate le condotte connesse al processo di acquisizione e utilizzo dei *Big Data*, che possono sollevare criticità di un qualche rilievo per l'*enforcement* sia delle norme a tutela della concorrenza che di quelle a tutela del consumatore.

5.4.1. La raccolta di dati

Il processo di raccolta dei dati è discusso approfonditamente nel capitolo 1 dell'Indagine, dove sono stati rilevati i *bias* comportamentali e le asimmetrie informative che caratterizzano il rapporto utente-operatore nonché la relazione (complessa) tra concorrenza e *privacy*. In tale contesto, sia gli strumenti di intervento di tutela del consumatore che quelli *antitrust* possono contribuire alla tutela degli utenti in questa fase delicata del rapporto con le piattaforme digitali.

Il contributo della tutela del consumatore. È ampiamente trattata negli studi economici più recenti che analizzano i *data-driven zero-price markets* la questione della potenziale efficacia degli strumenti di tutela del consumatore nel contribuire a risolvere le inefficienze derivanti dal valore pari a zero del prezzo e dalla mancanza di consapevolezza da parte degli utenti delle vicende che attengono ai dati che essi forniscono alle piattaforme digitali. Di rilievo, sotto tale profilo, sono le condizioni di fruizione dei servizi offerti gratuitamente all'interno di una piattaforma e, in senso più ampio, rispetto ai fornitori di servizi analoghi.

L'AGCM è più volte intervenuta con azioni di *enforcement* del Codice del Consumo nel sistema della "*data driven economy*", e specificatamente nei confronti di WhatsApp e Facebook, per tutelare i consumatori, soprattutto quelli fruitori di servizi digitali "pagati" con i dati personali.

L'Autorità ha, in particolare, ritenuto che i modelli di *business* incentrati sulla raccolta e l'elaborazione dei dati, anche quando l'utente riceve il servizio senza dover pagare un corrispettivo in termini monetari, rientrassero nella nozione di attività economica ai sensi del diritto europeo. A tal fine, l'Autorità, dando concreta attuazione a principi ormai consolidati sia a livello europeo che internazionale, ha ampliato la nozione di rapporto di consumo, riconoscendo la natura economica del comportamento dell'utente anche in relazione alle piattaforme digitali che offrono servizi gratuitamente.

Ciò posto, l'Autorità ha ritenuto ingannevole la schermata di registrazione ad un *social network* (Facebook) nella quale mancava un'adeguata e immediata informazione circa le finalità commerciali della raccolta dei dati dell'utente e ha ritenuto aggressive le modalità con cui il *social network* procedeva all'acquisizione del consenso per lo scambio, per fini commerciali, di dati dei propri utenti con siti *web* o *app* di terzi[225].

In un altro caso, l'Autorità ha ritenuto aggressiva la condotta di un fornitore di un servizio di messaggistica (WhatsApp) consistente nell'aver di fatto forzato i propri utenti ad accettare nuovi Termini di Utilizzo – relativi alle condizioni dei propri dati ai fini di profilazione commerciale e pubblicitari – facendo loro credere che sarebbe stato altrimenti impossibile proseguire nell'utilizzo dell'applicazione medesima[226].

La tutela del consumatore, dunque, può intervenire su una molteplicità di profili connessi al rapporto tra operatori e utenti nella fase di acquisizione dei dati. L'effetto utile di tale intervento non è solo quello di fornire una tutela diretta ai consumatori, ma anche quello di svolgere un ruolo pro-concorrenziale nella misura in cui gli utenti sono posti nella condizione di esercitare (più) consapevolmente e attivamente le proprie scelte di consumo: quanto più i consumatori sono informati, consapevoli e liberi nelle loro scelte, tanto più le imprese sono incentivate a competere tra di loro differenziando le proprie offerte di servizi digitali gratuiti in relazione alla qualità nella forma di *privacy*.

A questo riguardo, tenuto conto delle grandi dimensioni di molti operatori attivi nell'economia digitale e impregiudicato quanto già previsto dal RGPD, al fine di garantire un efficace effetto deterrente delle norme a tutela del consumatore, sarebbe necessario prevedere quanto prima un aumento del massimo edittale per le sanzioni, anche in linea con quanto già stabilito dalla recente direttiva (UE) 2019/2161 del Parlamento europeo e del Consiglio del 27 novembre 2019 che modifica

[225] Cfr. PS11112 - *Facebook-Condivisione dati con terzi*, 29 novembre 2018 n. 27432.
[226] Cfr. PS10601 - *Whatsapp-Trasferimento dati a Facebook*, 11 maggio 2017 n.26597.

la direttiva 93/13/CEE del Consiglio e le direttive 98/6/CE, 2005/29/CE e 2011/83/UE del Parlamento europeo e del Consiglio per una migliore applicazione e una modernizzazione delle norme dell'Unione relative alla protezione dei consumatori.

Gli spazi per gli interventi di tutela della concorrenza. La diretta applicabilità degli strumenti consumeristici non significa che questi siano gli unici disponibili, dal momento che essi si pongono in un rapporto di complementarietà con l'apparato di tutela della concorrenza e di quello a protezione dei dati personali. Da questo punto di vista, l'Autorità, accentrando due delle tre competenze (tutela della concorrenza e del consumatore), ha una maggiore flessibilità nella scelta dell'intervento più adatto alla specifica situazione posta sotto osservazione.

Il legame tra *Big Data*, *privacy* e *enforcement* della disciplina a tutela della concorrenza può interessare tutti gli strumenti di intervento dell'autorità *antitrust*: intese, abusi e concentrazioni.

Ad esempio, è possibile rilevare, in linea generale, come accordi orizzontali tra imprese che riducono il livello di *privacy* offerto sul mercato possano essere considerati restrittivi della concorrenza, al pari di accordi idonei ad aumentare i prezzi. In una prospettiva più "tradizionale", anche accordi tra imprese aventi ad oggetto la condivisione di dati personali dei propri utenti possono evidentemente presentare criticità laddove siano idonei ad agevolare un coordinamento delle politiche commerciali delle imprese stesse.

Con specifico riguardo allo scenario in cui il mercato è caratterizzato dalla presenza di un operatore in posizione dominante, si può porre la questione se l'acquisizione di "troppi" dati possa costituire un abuso di sfruttamento. Non si può escludere infatti, almeno in linea di principio, che talune modalità di acquisizione di dati personali possano integrare un abuso di posizione dominante di sfruttamento al pari della definizione di prezzi eccessivamente gravosi. Si tratta, peraltro, di una condotta che può potenzialmente anche presentare profili escludenti laddove l'utilizzo dei dati in questione sia funzionale all'adozione di condotte con un effetto di *foreclosure* in uno specifico mercato ovvero di estensione della dominanza in un mercato contiguo. Al contempo, occorre rilevare la complessità inerente la definizione di un *benchmark* di riferimento per valutare l'eccessività e iniquità dell'acquisizione di dati personali[227]. Ciò anche a causa del fatto che, ancor più che nella fattispecie "tradizionale" di prezzi ingiustificatamente gravosi (già di per sé assai problematica e soggetta a *standard* probatori particolarmente stringenti), appare poco utile il riferimento a un livello "competitivo" di *privacy*, idoneo a guidare le valutazioni di eccessività e iniquità. Inoltre, come rilevato nel primo capitolo dell'Indagine, la valutazione della sussistenza di un danno al consumatore può essere apprezzata solo considerando nel suo complesso la filiera dei *Big Data* che comprende non solo la loro raccolta, ma anche il loro utilizzo concreto.

Se il livello "competitivo" di *privacy*, per quanto illustrato nella sezione 5.3., non appare agevolmente identificabile (e in ogni caso potrebbe risultare socialmente insoddisfacente) un *benchmark* alternativo potrebbe rintracciarsi in quanto stabilito dalla regolazione - più attenta alle modalità con le quali vengono raccolti i dati personali che alla loro quantità - ovvero nel grado di genericità relativo

[227] Cfr. conclusioni dell'Avvocato Generale N. Wahl del 6 aprile 2017, nella causa C-177/16, *Biedrība «Autortiesību un komunicēšanās konsultāciju aģentūra – Latvijas Autoru apvienība» c. Konkurences padome*, in tema di determinazione della nozione di prezzi non equi utilizzata nell'articolo 102, co. 2, lett. a) del TFUE.

all'utilizzo di tali dati. Una genericità che può privare il singolo individuo della capacità di controllare la sorte delle informazioni che riguardano la sua persona e la sua dignità.

In entrambi i casi è ovvio che l'intreccio tra le competenze del Garante dei dati personali e dell'Autorità *antitrust* può rivelarsi inevitabile e particolarmente complesso. La cooperazione tra le due istituzioni, di cui questa Indagine è prova evidente, potrà contribuire, caso per caso, a selezionare la strumentazione più efficace.

Infine, la valenza economica dei dati personali può rilevare anche per la valutazione di merito svolta dalle autorità di concorrenza nell'ambito del controllo delle concentrazioni. In particolare – oltre alla rilevanza dei *Big Data* nella valutazione del potere di mercato delle imprese e delle barriere all'entrata – si pone la questione specifica dell'eventuale valutazione degli effetti di un'operazione di concentrazione sulla *privacy*, al pari dell'analisi su altri aspetti "tradizionali" quali i prezzi. Ciò appare particolarmente rilevante, ad esempio, nei casi in cui l'impresa acquisita adotta, a differenza dell'acquirente, un modello di *business* "a bassa intensità di dati" e vi è il rischio che l'operazione di concentrazione condizioni negativamente il livello di *privacy* disponibile nel mercato.

Ad esempio – sempre nell'ambito del quadro di regole definito dal RGPD – una riduzione qualitativa derivante da un accresciuto potere di mercato potrebbe concretizzarsi, tra l'altro, in un aumento del volume di dati richiesti, nell'utilizzo congiunto di più fonti di dati prima separate, nell'utilizzo di tali dati per un insieme più ampio di finalità ovvero nella riduzione del livello di controllo che gli utenti hanno sui propri dati.

Si tratta di aspetti che possono ben rientrare tra le valutazioni svolte da un'autorità di concorrenza, nella prospettiva di una interpretazione giustamente ariosa della nozione di benessere del consumatore che costituisce obiettivo fondamentale anche del controllo delle concentrazioni.

5.4.2. L'utilizzo dei *Big Data* per la personalizzazione dei servizi

I *Big Data* sono spesso utilizzati dalle imprese per offrire servizi altamente personalizzati agli utenti finali. Ad esempio, piattaforme di *e-commerce* e piattaforme attive nella distribuzione di contenuti digitali possono utilizzare i *Big Data* per offrire agli utenti suggerimenti di prodotti e servizi da acquistare o di contenuti editoriali e audiovisivi da fruire. Altre piattaforme come i motori di ricerca e i *social network* possono personalizzare i risultati delle ricerche e i contenuti presentati all'attenzione degli utenti finali sulla base delle informazioni che hanno raccolto ed elaborato sui singoli individui[228].

I sistemi *software* su *web* adattano la loro configurazione e il loro funzionamento a ogni singolo utente o gruppi di utenti e implementano diversi tipi di personalizzazione. Ad esempio: *i)* i sistemi di ricerca adattivi nelle piattaforme promuovono nei risultati i contenuti che ritengono più pertinenti per l'utente in base ai suoi interessi e ai suoi bisogni, così come si sono manifestati attraverso le precedenti esperienze di consumo; *ii)* i sistemi di adattamento delle pagine producono su misura il contenuto di una pagina *web* in base al profilo dell'utente portando i *link* consigliati in primo piano; *iii)* i filtri e i sistemi di raccomandazione completano la ricerca e la navigazione in base alle informazioni personali, suggerendo attivamente articoli che sembrano più rilevanti per l'interesse dell'utente.

La personalizzazione implica una profilazione degli utenti basata su dati personali, cioè su informazioni riferite o riferibili a soggetti identificati e consiste nell'utilizzo di tali dati per valutare

[228] Cfr. audizione dei Proff. Preta, Maggiolino e altri (30 gennaio 2018).

determinati aspetti relativi a uno specifico individuo, in particolare per analizzare o prevedere aspetti riguardanti il rendimento professionale, la situazione economica, la salute, le preferenze personali, gli interessi, l'affidabilità, il comportamento, l'ubicazione o gli spostamenti di detta persona fisica. Al fine di creare i profili degli utenti e di personalizzare le pagine web possono essere impiegate tecniche di *machine learning* e di apprendimento automatico[229]. Tale tecniche, ad esempio, sono utilizzate comunemente sulla cronologia di navigazione, sui log dei motori di ricerca, sui dati personali immessi dall'utente, sul comportamento dell'utente mentre interagisce con il sistema e sui dati geografici.

Dal punto di vista degli utenti, la personalizzazione può determinare anche notevoli benefici, ad esempio in termini di riduzione di costi di ricerca e di transazione. In alcuni casi, nelle piattaforme multi-versante, il vantaggio della personalizzazione va a favore non solo dell'utente profilato, ma anche e soprattutto degli operatori pubblicitari *online* che possono così meglio indirizzare i loro messaggi[230]. Peraltro, considerando che nei c.d. mercati senza prezzo l'utente paga con i propri dati individuali e con l'attenzione che rivolge ai messaggi pubblicitari veicolati dalla piattaforma, una pubblicità altamente personalizzata consente di limitare l'esposizione a messaggi privi di interesse per l'utente.

La personalizzazione del servizio offerto all'utente non crea, di per sé, preoccupazioni concorrenziali mentre invece presenta importanti profili di criticità sotto il profilo della protezione dei dati. Tuttavia, potenziali rischi concorrenziali potrebbero emergere in quelle situazioni in cui la personalizzazione viene effettuata da una piattaforma dominante che opera sia come "intermediario" che come fornitore all'utente del servizio oggetto di intermediazione (in concorrenza con altri operatori). In tali circostanze, infatti, la personalizzazione dei risultati di ricerca potrebbe rendere più agevole la realizzazione di pratiche di *search discrimination*, che un'impresa potrebbe porre in essere al fine di favorire i propri prodotti e servizi venduti sulla piattaforma.

Personalizzazione dei servizi e tutela del consumatore. La personalizzazione dei servizi offerti può generare potenziali criticità sotto il profilo della tutela del consumatore, che potrebbe non essere posto nelle condizioni di sapere in che misura il servizio di ricerca/intermediazione utilizzato filtra o adegua i risultati sulla base delle caratteristiche individuali dell'utente.

L'asimmetria informativa tipica delle piattaforme digitali comporta una scarsissima consapevolezza degli utenti circa l'effettiva portata e pervasività della raccolta e dell'utilizzo dei propri dati a fini commerciali da parte delle imprese operanti *on line* (cfr. i risultati della menzionata *survey* condotta dall'Autorità nel mese di febbraio 2018).

La scarsa consapevolezza in merito alla raccolta dei dati e alle relative modalità di utilizzo rende il consumatore particolarmente vulnerabile e privo di reale autodeterminazione nel momento in cui è di fronte alla scelta di avvalersi o meno di servizi *on line*, ancor più nell'ipotesi in cui tali servizi siano pubblicizzati come gratuiti e i dati forniti dall'utente costituiscano il solo valore di scambio.

In particolare, l'utente viene attratto dal *claim* sulla gratuità del servizio offerto nonché dall'agevole e immediata fruibilità dello stesso per poi essere indotto a condividere in rete esperienze personali, dati e informazioni, nella convinzione che tale attività sia esclusivamente funzionale alla stessa

[229] Cfr. OCSE (2018), *Personalised Prices in the Digital Era*, Parigi, novembre.
[230] Cfr. audizione dei Proff Giannotti e Pedreschi (5 dicembre 2017) e audizioni di Mediaset (28 novembre 2017) e Facebook (5 febbraio 2018).

operatività e finalità del servizio fornito. Non sono invece chiaramente rese note le finalità remunerative perseguite dalle imprese, che possono ottenere il massimo sfruttamento economico dei dati attraverso la profilazione degli utenti. Una sempre più accurata profilazione dell'utenza consente infatti agli operatori di raggiungere target specifici di consumatori, indirizzando loro messaggi mirati, con crescenti livelli di personalizzazione.

La tutela del consumatore può senz'altro intervenire al fine di garantire che gli utenti siano posti nella condizione di esercitare consapevolmente e attivamente le proprie scelte di consumo. Ad esempio, nel citato provvedimento adottato dall'Autorità contro Facebook (2018), è stata ritenuta ingannevole la schermata di registrazione al *social network* nella quale mancava un'adeguata e immediata informazione circa le finalità commerciali della raccolta dei dati dell'utente. Le informazioni fornite risultavano infatti generiche ed incomplete e non distinguevano adeguatamente tra, da un lato, l'utilizzo dei dati funzionale alla personalizzazione del servizio con l'obiettivo di facilitare la socializzazione con altri utenti "consumatori", e, dall'altro, l'utilizzo dei dati per realizzare campagne pubblicitarie mirate.

Oltre ai citati aspetti informativi, si pone una questione più generale di benessere dei consumatori. L'analisi dei dati individuali e la conseguente personalizzazione dell'offerta può infatti tradursi in una oggettiva difficoltà per gli utenti di confrontare le caratteristiche e i contenuti di offerte diverse, ostacolando così la comparabilità dei prodotti e la stessa libertà di scelta, con ricadute negative anche in termini concorrenziali. La personalizzazione e differenziazione dei prodotti offerti può infatti portare a forme di "offuscamento" facilitando peraltro, per questa via, equilibri collusivi caratterizzati da prezzi più elevati.

Personalizzazione dei servizi e pluralismo. Le preoccupazioni maggiori connesse al fenomeno della personalizzazione dei servizi, tuttavia, risiedono probabilmente in relazione alla tutela del pluralismo.

Sebbene, infatti, i singoli utenti possono potenzialmente beneficiare dell'accesso ai contenuti giornalistici di maggiore interesse, un'elevata personalizzazione nella distribuzione di contenuti giornalistici può ridurre significativamente il grado di pluralismo nell'informazione e, conseguentemente, la possibilità per il consumatore di accedere a una pluralità e varietà di fonti informative. Su questo tema si rinvia a quanto illustrato *supra* (§ 3.1).

5.4.3. L'utilizzo dei *Big Data* per la personalizzazione dei prezzi

L'avvento dei *Big Data* consente sempre più alle imprese di raccogliere dati personali dei consumatori e di utilizzare gli algoritmi per attuare forme avanzate di discriminazione di prezzo[231].

I mercati in cui è possibile ipotizzare un maggiore sviluppo di strategie di personalizzazione dei prezzi sono quelli caratterizzati:

- da un elevato grado di potere di mercato;
- da una maggiore possibilità di identificare la disponibilità a pagare dei consumatori, particolarmente agevole per talune piattaforme e operatori *online*;
- dall'assenza della possibilità di arbitraggio tra i consumatori con una bassa ed elevata disponibilità a pagare; questo è riscontrabile in alcuni servizi, acquistabili *online*, che sono tipicamente non trasferibili - quali prenotazioni alberghiere, biglietti aerei, per concerti, per

[231] OCSE (2018), *Personalised Prices in the Digital Era*, Parigi, novembre.

musei - e per contenuti digitali accessibili solo da un dispositivo o da un conto personale - quali film, *e-book*, abbonamenti a giornali, mentre è più difficile con riguardo a beni fisici durevoli, quali c*omputer* e vestiti.

Prezzi personalizzati, efficienza e aspetti distributivi. In generale, prezzi personalizzati possono migliorare l'efficienza allocativa, statica e dinamica[232]:

- dal punto di vista statico, i prezzi personalizzati hanno il potenziale di migliorare, attraverso un aumento delle quantità scambiate, il benessere sociale;
- i prezzi personalizzati possono migliorare l'efficienza dinamica, incrementando gli incentivi all'innovazione per via dei profitti addizionali che possono essere acquisiti dall'impresa.

Sotto una diversa prospettiva, l'applicazione di prezzi personalizzati può avere anche implicazioni di carattere distributivo, che possono interessare diverse (categorie di) soggetti. Infatti, se da un lato la discriminazione di prezzo può consentire un aumento dei profitti delle imprese a danno del benessere dei consumatori, dall'altro lato può anche avere un impatto eterogeneo su diversi gruppi di consumatori, aumentando il benessere di taluni e riducendo il benessere di altri.

In quest'ultimo scenario, anche in una prospettiva di *enforcement* fondata sulla tutela del benessere dei consumatori, diventano particolarmente complessi interventi sul tema dei prezzi personalizzati, dal momento che potrebbero comportare la necessità di confrontare la posizione di gruppi diversi di consumatori.

La tutela della concorrenza. La normativa europea a tutela della concorrenza vieta *esplicitamente* solo le discriminazioni di prezzo nei confronti di imprese (in ragione delle loro implicazioni escludenti) e non anche dei consumatori finali. Non è, pertanto, del tutto chiaro fino a che punto le disposizioni relative a condotte discriminatorie si applichino anche ai rapporti *business-to-consumer*. In questo contesto, l'art. 102.c TFUE è stato, ad oggi, per lo più utilizzato con riguardo a questioni legate alla protezione del mercato interno concernenti discriminazioni basate sul paese di residenza dei clienti e mai a prezzi personalizzati. Alternativamente, ci si potrebbe chiedere se simili pratiche possano essere perseguite nell'ambito dell'art. 102.a TFUE, dovendo però chiarire quale sia il *test* che andrebbe applicato in siffatta ipotesi. In particolare, appare particolarmente complesso ipotizzare condotte di sfruttamento quando i prezzi personalizzati hanno un impatto negativo sul benessere di alcuni consumatori, quelli con la disponibilità a pagare maggiore, ma hanno un impatto positivo sul benessere di altri consumatori, quelli con la disponibilità a pagare minore.

La tutela del consumatore. La tutela del consumatore può avere già oggi un ruolo prominente nel trattare i rischi derivanti dai prezzi personalizzati nell'ambito delle pratiche commerciali scorrette.

In primo luogo, potrebbe essere valutato se configura una pratica commerciale scorretta l'applicazione di prezzi personalizzati in maniera non trasparente, o senza fornire ai consumatori la possibilità di *opt-out*. Ciò di fatto obbligherebbe le imprese a fornire informazioni con riguardo alle proprie strategie di prezzo in modo tale da consentire ai consumatori di acquisire consapevolezza dell'esistenza di tali pratiche e adottare, ove necessario, specifiche azioni per eluderle. Le evidenze disponibili appaiono confermare che la reazione dei consumatori può differire laddove siano a

[232] Cfr. OCSE (2016), *Price Discrimination*, Background note by the Secretariat, Parigi, novembre; McSweeney, T. e O Dea B., *The Implications of Algorithmic Pricing for Coordinated Effects Analysis and Price Discrimination Markets in Antitrust Enforcement*, Antitrust, Vol. 32, No. 1, Fall 2017.

conoscenza della raccolta dati e della personalizzazione così come quando hanno la possibilità di *opt-out*[233].

In secondo luogo, si potrebbero anche perseguire e sanzionare pratiche commerciali scorrette ancillari idonee a re-inforzare gli effetti negativi dei prezzi personalizzati, come pratiche ingannevoli o omissive che limitano ulteriormente la trasparenza e la scelta del consumatore[234].

5.4.4. Condotte che possono integrare possibili abusi di posizione dominante

I casi della Commissione europea. La repressione degli abusi di natura escludente costituisce una priorità nell'*enforcement* dell'art. 102 TFUE e le recenti esperienze applicative della Commissione europea mostrano come tale norma sia idonea a contrastare diverse pratiche escludenti connesse all'utilizzo dei *Big Data*, volte a frapporre ostacoli ai soggetti terzi nell'acquisizione dei dati degli utenti o consistenti in pratiche discriminatorie o leganti.

La Commissione europea ha concluso ben tre procedimenti istruttori nei confronti di Google per abusi di posizione dominante aventi ad oggetto condotte escludenti nei mercati delle ricerche generiche su Internet e in quello adiacente dell'intermediazione pubblicitaria nei motori di ricerca. Tali pratiche hanno permesso all'*incumbent* di rafforzare la propria posizione dominante nei mercati interessati e di acquisire una mole sempre più significativa di dati degli utenti, preziosi per le sue attività di ricerca e di pubblicità *on line*.

In un caso recente (marzo 2019), è stata comminata a Google una sanzione pari a 1,49 miliardi di euro, per aver abusato della propria posizione dominante sul mercato dell'intermediazione pubblicitaria nei motori di ricerca, dove Google è attiva tramite la piattaforma *AdSense for Search*. Attraverso *AdSense for Search* Google agisce come un intermediario pubblicitario, tra inserzionisti e proprietari di siti *web* che intendono trarre profitto dallo spazio attorno alle pagine dei risultati della ricerca[235]. Dall'istruttoria condotta dalla Commissione europea è emerso che Google ha dapprima imposto ai principali siti *publisher* un obbligo di fornitura esclusiva, che impediva ai concorrenti di inserire annunci pubblicitari collegati alle ricerche sui siti *web* più significativi dal punto di vista commerciale, e ha poi adottato una strategia di "esclusiva non rigida", volta a riservare gli spazi migliori per i propri annunci collegati alla ricerca e a controllare le prestazioni degli annunci dei concorrenti[236].

---

[233] European Commission, "*Market study on online market segmentation through personalised pricing/offers in the EU*", 2018, https://ec.europa.eu/info/publications/consumer-market-study-online-market-segmentation-through-personalised-pricing-offers-european-union_en.

[234] Ad esempio, affermare che un prezzo è il più conveniente mentre ad altri consumatori vengano offerti prezzi migliori, offrire un prezzo personalizzato scontato che è più elevato di quello pubblico, raccogliere dati per personalizzare i prezzi senza il consenso dei consumatori.

[235] Quando un utente effettua una ricerca utilizzando questa funzione, insieme ai risultati della ricerca, il sito web propone annunci pubblicitari collegati alla ricerca. Dal momento che i concorrenti nella pubblicità collegata alle ricerche, come Microsoft e Yahoo, non hanno la possibilità di vendere spazi pubblicitari nelle pagine dei risultati di ricerca di Google, i siti web di terzi rappresentano un importante punto di accesso per tentare di competere efficacemente con Google.

[236] Commissione Europea (2019), "*AT.40411 Google Search (AdSense)*", cfr. http://europa.eu/rapid/press-release_IP-19-1770_it.htm. In particolare, la Commissione ha accertato che:

- dal 2006 molti contratti tra Google e i siti *publisher* più rilevanti prevedevano clausole di esclusiva (i *publisher* avevano il **divieto di mostrare sulle pagine dei risultati di ricerca annunci pubblicitari collegati alla ricerca dei concorrenti)**;
- a partire dal marzo 2009 Google aveva gradualmente iniziato a sostituire le clausole di esclusiva con le cosiddette clausole di "posizionamento premium", che imponevano ai *publisher* di riservare lo spazio più redditizio sulle pagine dei risultati di ricerca agli annunci di Google e di prevedere un numero minimo di annunci di Google;

Nel luglio 2018 la Commissione europea ha inflitto a Google un'altra ammenda pari a 4,34 miliardi di euro per abuso di posizione dominante, accertando che l'operatore, consapevole del rilevante passaggio dall'utilizzo dei computer *desktop* a quello di Internet mobile, aveva implementato una strategia per far sì che gli utenti, anche in tale fase di transizione, continuassero ad usare *Google Search* sui propri dispositivi mobili. Google aveva infatti imposto ai produttori di dispositivi Android e agli operatori di reti mobili condizioni contrattuali illegittime, al fine di consolidare la propria posizione dominante nel mercato delle ricerche generiche su Internet. Tali condotte avevano consentito a che il motore di ricerca e il *browser* di Google venissero preinstallati sulla quasi totalità dei dispositivi Android di Google, a scapito dei motori di ricerca concorrenti[237].

Infine, nel giugno 2017 la Commissione ha accertato che Google aveva abusato della propria posizione dominante nel mercato dei servizi di ricerca generica, riservando un trattamento più favorevole, in termini di posizionamento e di visualizzazione nelle sue pagine generali dei risultati di ricerca, al proprio servizio di acquisti comparativi rispetto ai servizi concorrenti. In particolare, il servizio di Google non era soggetto agli algoritmi specifici che rendevano probabile la retrocessione dei servizi di acquisti comparativi concorrenti all'interno delle pagine di ricerca generica di Google. Inoltre, il servizio di acquisti comparativi di Google veniva visualizzato con funzionalità migliorate in cima ai risultati della prima pagina di ricerca generica, o comunque tra i primi risultati, mentre tali funzionalità non erano accessibili ai concorrenti. La Commissione ha dunque comminato un'altra ammenda pari a 2,42 miliardi di euro[238].

Le tre importanti decisioni della Commissione europea riguardano casi che rientrano a pieno titolo nel solco delle condotte escludenti tradizionalmente contestate dalla disciplina *antitrust*, opportunamente interpretate alla luce delle specificità dell'economia digitale. Anche nel caso più risalente, la decisione del 2017 relativa ai servizi di ricerca generica, che vedeva l'utilizzo dell'algoritmo come principale strumento di discriminazione, l'accertamento dell'abuso si fondava soprattutto sulla portata escludente della strategia di Google.

In questa prospettiva, rileva pertanto sottolineare come, anche nei mercati *data driven*, sembrino riproporsi condotte restrittive riconducibili a fattispecie familiari all'*enforcement antitrust*, che senza dubbio dispone di strumenti sufficientemente flessibili per un'applicazione evolutiva delle norme.

Possibili strategie abusive. Al di là delle istruttorie sopra descritte che hanno portato all'accertamento di diversi abusi di posizione dominante, l'utilizzo di *Big Data* può potenzialmente

---

- dal marzo 2009 Google aveva altresì previsto clausole che imponevano ai *publisher* di ottenere l'autorizzazione scritta di Google prima di modificare le modalità di visualizzazione dei messaggi pubblicitari dei concorrenti.

[237] Commissione Europea (2018), "*AT.40099 Google Android*", http://ec.europa.eu/competition/antitrust/cases/dec_docs/40099/40099_9993_3.pdf
In parrticolare, Google aveva:

- imposto ai produttori di dispositivi mobili intelligenti di preinstallare l'applicazione *Google Search* e la sua applicazione di *browsing* (*Chrome*) come condizione per la concessione della licenza relativa al portale di vendita di applicazioni di Google (*Play Store*);
- concesso significativi incentivi finanziari ad alcuni grandi produttori e operatori di reti mobili affinché preinstallassero a titolo esclusivo l'applicazione *Google Search* sui loro dispositivi e
- impedito ai produttori che desiderassero preinstallare le applicazioni Google la vendita di dispositivi mobili funzionanti con versioni alternative di *Android* non approvate da Google (le cosiddette "*Android forks*").

[238] Commissione Europea (2017), "*AT.39740 Google Search (Shopping)*", http://ec.europa.eu/competition/antitrust/cases/dec_docs/39740/39740_14996_3.pdf.

sollevare preoccupazioni concorrenziali – la cui trattazione non può che essere di natura meramente esemplificativa – in diversi snodi dell'ecosistema.

In via preliminare si può comunque osservare come nell'economia digitale la definizione dei mercati rilevanti e l'accertamento del potere di mercato siano indubbiamente più complessi che nell'economia tradizionale. Ciò posto, impregiudicata l'utilità di una comprensione del contesto in cui le condotte oggetto di analisi si sviluppano e producono i loro effetti, una maggiore attenzione può essere prestata direttamente alla portata escludente di tali condotte, in particolare se fondate sulla centralità e non replicabilità dei dati nella disponibilità dell'impresa dominante che possono interessare contemporaneamente una varietà di mercati.

Abusi di sfruttamento. Eventuali abusi di sfruttamento possono originare dall'incontrastato potere di mercato che alcuni operatori detengono nei c.d. mercati "senza prezzo" e dalle modalità con le quali vengono raccolti i dati individuali.

Sebbene gli abusi di sfruttamento costituiscano una dimensione residuale dell'*enforcement antitrust* "tradizionale", il loro rilievo nei mercati digitali appare potenzialmente più esteso. Oltre al tema connesso al rapporto tra piattaforme e utenti finali con specifico riferimento al trattamento dei dati personali e, dunque, alla possibilità di configurare possibili abusi di posizione dominante anche con riferimento a tale aspetto (cfr. *supra*), il potere di mercato degli operatori digitali può essere esercitato anche attraverso l'imposizione di prezzi (o altre condizioni contrattuali) eccessivamente gravosi. L'esistenza di posizioni dominanti nell'attività di intermediazione tra una pluralità di soggetti fa sì che le preoccupazioni possano riguardare uno solo dei versanti della piattaforma e, dunque, una sola categoria di utenti. Ad esempio, alcune recenti iniziative (legislative e di *enforcement*) pongono l'accento sulla relazione tra le piattaforme di intermediazione nel commercio elettronico e gli utenti non consumatori di tali piattaforme. Si tratta, dunque, di fattispecie il cui trattamento richiede una chiara definizione degli obiettivi che l'*enforcement* persegue, soprattutto in considerazione del fatto che tali iniziative possono comportare la necessità di un bilanciamento tra il benessere degli utenti dei diversi versanti della piattaforma, ovvero più direttamente tra i soggetti imprenditoriali che si avvalgono delle piattaforme e i consumatori finali.

Dati e rifiuto a contrarre. L'articolo 102 TFUE trova applicazione nei casi in cui il detentore di un'*essential facility* opponga un rifiuto a contrarre a un'impresa con la quale compete in un mercato a valle. Anche nel caso in cui l'*essential facility* sia costituita da dati, un eventuale rifiuto a concedere a terzi l'accesso a tali dati ha una rilevanza *antitrust* se e nella misura in cui è idoneo a ridurre la concorrenza in un mercato complementare/a valle. Pertanto, ai fini dell'applicazione dell'art. 102 TFUE, assume particolare peso la *finalità* alla base di una richiesta di accesso ai dati detenuti da un'impresa dominante. Le richieste di accesso ai dati potenzialmente più rilevanti in una prospettiva concorrenziale sono quelle relative ai dati: *i)* necessari per offrire un bene/servizio al consumatore nel mercato in cui i dati sono acquisiti, in concorrenza con l'operatore (dominante); ovvero; *ii)* necessari per competere in un mercato contiguo; o *iii)* in un *aftermarket* in cui è attivo l'operatore in posizione dominante.

Ai fini dell'analisi dell'indispensabilità tipica della dottrina *antitrust* dell'*essential facility* nel settore dei *Big Data*, almeno tre aspetti specifici appaiono potenzialmente rilevanti:

- la natura personale o meno dei dati oggetto della richiesta di accesso;

- se i dati in questione siano stati: *i)* volontariamente forniti dal soggetto a cui si riferiscono; *ii)* rilevati dall'operatore dominante; *iii)* ricavati tramite attività di analisi dei dati svolte dall'operatore in questione (*analytics*);
- il grado di aggregazione dei dati oggetto della richiesta di accesso potendo distinguere, dunque, tra dati a livello individuale, aggregati o *bundled*.

In ogni caso, la specificità, la quantità e la qualità dei dati possono configurare un ostacolo alla concorrenza e favorire una condotta abusiva, nella forma di un rifiuto a contrarre, solo laddove tali dati integrino i requisiti stringenti di una *essential facility* per la fornitura di un particolare servizio.

Solo in circostanze eccezionali, dunque, i *Big Data* raccolti da un'impresa possono costituire una risorsa "essenziale" per operare in un mercato ed essere soggetti ad un obbligo a contrarre ai sensi della normativa a tutela della concorrenza. La nozione legale di *essential facility* va oltre il mero riconoscimento della rilevanza dei *Big Data* nel processo competitivo. Infatti, anche quando una risorsa è un'importante fonte di vantaggio competitivo e costituisce una barriera all'entrata, la normativa *antitrust* non impone necessariamente alle imprese di condividere tale risorsa con i propri concorrenti. Come è noto, un rifiuto a contrarre costituisce una violazione della normativa *antitrust* se ricorrono le seguenti condizioni cumulative: *i)* il rifiuto si riferisce ad un prodotto o ad un servizio obiettivamente necessario per poter competere in maniera effettiva su un mercato a valle, *ii)* è probabile che il rifiuto determini l'eliminazione di una concorrenza effettiva sul mercato a valle, e *iii)* è probabile che il rifiuto determini un danno per i consumatori. Quando il rifiuto a contrarre concerne l'esercizio di diritti di proprietà intellettuale, la giurisprudenza euro-unitaria ha aggiunto la condizione ulteriore che il rifiuto si debba riferire all'offerta di un prodotto o servizio nuovo per il quale sussiste una potenziale domanda[239].

La particolare cautela che la giurisprudenza impone nell'esame delle condotte di rifiuto a contrarre è giustificata dall'esigenza di tutelare tanto la concorrenza nella fornitura dei servizi realizzati tramite l'utilizzo dei *Big Data*, quanto la concorrenza nelle attività di raccolta e analisi dei dati, che possono portare benefici ai consumatori nella forma di servizi innovativi, spesso a prezzi monetari nulli.

Se, da un lato, la disponibilità di dati costituisce un *asset* sempre più necessario per operare in una varietà di mercati, dall'altro lato, dati utili alla fornitura di un particolare servizio digitale possono essere raccolti da una varietà di fonti, a costi non necessariamente elevati (anche perché non vi è rivalità nel consumo dei dati, considerato che gli utenti possono condividere i medesimi propri dati con più imprese). Inoltre, viene spesso rilevato come non è tanto la raccolta dei dati in sé, quanto la capacità di estrarre informazioni utili da grandi volumi e varietà dei dati a rappresentare la vera risorsa scarsa alla base di rilevanti posizioni dominanti[240].

Occorre, peraltro, considerare anche la relazione tra un eventuale obbligo a fornire dati e il RGPD, almeno sotto due aspetti:

- in primo luogo, occorre tenere in considerazione il diritto alla portabilità dei dati personali previsto dal RGPD, posto che tale portabilità può potenzialmente consentire a un'impresa di acquisire i dati di interesse direttamente dai soggetti a cui si riferiscono senza che sia

[239] Corte di Giustizia dell'Unione Europea, C-241/91 P e C-242/91, *Radio Telefis Eireann* (RTE) e *Independent Television Publications (ITP)/ Commission (Magill)*; C-418/01, *IMS Health/NDC Health*; T-201/04, *Microsoft/Commission*.
[240] OCSE (2016), "*Big data: bringing competition policy to the digital era - Background note by the Secretariat*", https://one.oecd.org/document/DAF/COMP(2016)14/en/pdf, pag. 22.

necessaria la fornitura dell'accesso ai dati da parte dell'impresa in posizione dominante. D'altro canto, come è noto, il RGPD non prevede un diritto alla portabilità di qualsiasi tipo di dati personali né un diritto alla portabilità continua e potenzialmente in tempo reale dei propri dati;

- in secondo luogo, occorre considerare la possibile tensione che può venirsi a determinare tra l'accesso ai dati e il diritto alla protezione dei dati e alla *privacy* dei soggetti ai quali i dati si riferiscono. Tale tensione può essere potenzialmente risolta attraverso tecniche di anonimizzazione dei dati o ai sensi di quanto previsto dall'art. 6 del RGPD.

In ogni caso, anche in assenza della configurazione di una *essential facility* ai sensi della disciplina a tutela della concorrenza, eventuali necessità di accedere e condividere determinate categorie di dati possono essere legittime per finalità diverse da quelle *antitrust*, ad esempio per garantire la salute pubblica. Si tratta di ipotesi che dovrebbero essere limitate ai casi in cui ciò sia strettamente necessario e proporzionato rispetto a rilevanti interessi pubblici primari, per le quali lo strumento più appropriato può essere quello della regolazione[241]. Ad esempio, regolazioni settoriali che consentano allo Stato di accedere a banche dati raccolte da imprese private e utili per ragioni di salute pubblica, ambientali, sicurezza, mobilità, sembrano lo strumento più appropriato per garantire obiettivi di interesse pubblico ed evitare inutili e costose duplicazioni di dati già disponibili.

Condotte escludenti. Più in generale, le attività di analisi ed elaborazione dei dati (*analytics*, *cloud computing*, *data storage*) possono favorire l'attuazione di condotte escludenti potenzialmente più diffuse. La capacità e gli incentivi delle imprese di porre in essere condotte restrittive è peraltro influenzata dall'elevato grado di integrazione verticale e conglomerale che caratterizza l'ecosistema digitale. Ad esempio, possono essere individuate due situazioni particolarmente delicate. La prima è quella in cui un operatore dominante fornisce alle imprese (e/o ai consumatori) una pluralità di servizi complementari, potendo ad esempio attuare pratiche leganti idonee a proteggere o estendere la propria posizione dominante. La seconda è quella tipica dell'integrazione verticale, in cui l'operatore dominante fornisce un servizio all'impresa terza con la quale compete in un diverso livello della filiera, situazione che potrebbe favorire l'emergere di condotte discriminatorie di natura escludente.

- ✓ *Leverage della posizione dominante*. Possibili abusi di posizione dominante possono avere luogo quando un'impresa in posizione dominante utilizza i dati raccolti in un mercato per estendere indebitamente il proprio potere di mercato attraverso condotte anti-competitive, quali vendite abbinate. Si tratta di un'ipotesi che impone un'attenta analisi ai fini di distinguere quelle condotte che possono in realtà avere un effetto pro-competitivo da quelle effettivamente idonee ad avere un effetto escludente pregiudizievole per la concorrenza e i consumatori.
- ✓ *Condotte discriminatorie*. Alla luce del potere di mercato che le maggiori piattaforme *online* detengono in attività di intermediazione di grande rilievo economico (e sociale), possono assumere una particolare valenza anche le possibili condotte escludenti poste in essere in mercati *data-driven* con un connotato discriminatorio. Alcune condotte di natura discriminatoria potenzialmente anti-concorrenziale alle quali appare opportuno prestare particolare attenzione sono quelle poste in essere da un operatore in posizione dominante che

[241] Cfr. Drexl J. (2016), Designing Competitive Markets for Industrial Data - Between Propertisation and Access, Max Planck Institute for Innovation & Competition Research Paper No. 16-13

svolga un'attività di intermediazione e al tempo stesso sia attivo come "utente" in (almeno) uno dei versanti della piattaforma in questione. Negli ecosistemi digitali i rapporti che legano gli operatori che detengono le piattaforme e gli operatori che utilizzano, o sono comunque soggetti all'attività di intermediazione svolta da tali piattaforme, sono particolarmente complessi. Tuttavia, per certi versi, i rischi che emergono sono analoghi a quelli che possono emergere in filiere di mercato "tradizionali" in cui un operatore in posizione dominante che eroga un servizio "essenziale" per l'attività a valle è verticalmente integrato.

- ✓ *Reducing rivals' data*. Nei mercati in cui la disponibilità di *Big Data* costituisce un'importante fonte di vantaggio competitivo, condotte che hanno un effetto di preclusione anticoncorrenziale possono essere realizzate tramite strategie che potrebbero essere definite come "*reducing rivals' data*". Ad esempio, condotte abusive di natura escludente possono essere poste in essere da un operatore in posizione dominante che impedisce ai propri concorrenti di accedere ai dati a causa di vincoli contrattuali imposti per l'utilizzo di determinati servizi, degli accordi di esclusiva stipulati con soggetti terzi o attraverso la creazione di ostacoli per l'utilizzo da parte dei consumatori dei servizi offerti da operatori concorrenti che consentirebbero a questi ultimi di acquisire dati rilevanti ad operare sul mercato. Si tratta di condotte il cui apprezzamento può richiedere un'analisi complessa, in cui l'esercizio della posizione dominante in un mercato può spesso avere effetti pregiudizievoli sulla concorrenza in altri mercati contigui.

5.4.5. L'utilizzo di *Big Data*, algoritmi di prezzo e collusione online

L'utilizzo dei *Big Data* può avere un impatto sul livello dei prezzi anche per il tramite dell'utilizzo di algoritmi di *pricing*. Un algoritmo di *pricing* è una procedura automatizzata usata per determinare i prezzi di vendita ottimali di prodotti/servizi sulla base delle condizioni del mercato e adeguarli in "tempo reale" alle variazioni di queste ultime.

Gli algoritmi di *pricing* possono essere implementati dalle stesse imprese che vendono il prodotto/servizio o da *software house* che offrono soluzioni complete per una gestione automatizzata della definizione dei prezzi. L'esistenza di prodotti di questo tipo fa sì che algoritmi di *pricing* anche molto sofisticati possano essere utilizzati da imprese di dimensione relativamente piccola.

Gli algoritmi di *pricing* usano largamente sistemi di monitoraggio, i quali acquisiscono i prezzi applicati dalle imprese concorrenti e sulla base di questi ricalcolano e aggiornano frequentemente i prezzi. I dati elaborati da un algoritmo di *pricing* al fine di computare il prezzo di un bene o di un servizio includono: i costi dell'impresa, i dati storici relativi a prezzi e profitti, i prezzi dei concorrenti, le informazioni personali del consumatore, ecc.

Gli algoritmi di *pricing* possono anche essere molto semplici e basarsi su regole predefinite, come la regola di allinearsi al prezzo più basso del mercato o di rimanere al di sotto/sopra di una determinata soglia rispetto al prezzo più basso del "mercato" di riferimento. Algoritmi di *pricing* più avanzati possono essere basati su modelli predittivi e sull'utilizzo di tecniche di *machine learning*, che "imparano" autonomamente le strategie ottimali di prezzo al fine di massimizzare i profitti dell'impresa.

L'elevata trasparenza dei mercati *online*, ossia l'ampia disponibilità di dati sui prezzi dei concorrenti e di altre informazioni rilevanti, la frequenza di aggiustamento dei prezzo, ossia la capacità degli algoritmi di monitorare in tempo reale i mercati potendo modificare istantaneamente e continuamente

i prezzi, nonché le capacità di apprendimento delle strategie di prezzo ottimali attraverso il *machine learning* fanno sì che l'utilizzo degli algoritmi possa potenzialmente agevolare fenomeni collusivi, più o meno taciti, e dunque più o meno leciti.

La difficoltà di rintracciare l'ingrediente decisivo per una violazione dell'art. 101 TFUE – lo scambio di volontà – in presenza di algoritmi sofisticati, caratterizzati da meccanismi di *machine learning* è, a dir poco, complicata.

LINEE GUIDA E RACCOMANDAZIONI DI POLICY

1. Governo e Parlamento si interroghino sulla necessità di promuovere un appropriato quadro normativo che affronti la questione della piena ed effettiva trasparenza nell'uso delle informazioni personali (nei confronti dei singoli e della collettività).

L'utilizzo intensivo dei *Big Data* costituisce un fenomeno che interessa sempre più l'intera economia e società. Agli indubbi vantaggi in termini di riduzione dei costi di transazione per imprese e cittadiniconsumatori, si affiancano nuovi rischi sotto il profilo concorrenziale, della protezione del dato personale e del pluralismo informativo.

In particolare, la disponibilità in capo ai grandi operatori digitali, attivi su scala globale, di enormi volumi e varietà di dati (personali e non personali, strutturati e non strutturati) e della capacità di analizzarli ed elaborarli ha dato luogo a inedite forme di sfruttamento economico del dato e della sua

valorizzazione ai fini della profilazione algoritmica legata a diversi scopi commerciali, generando nuove concentrazioni di potere, inteso non solo come 'potere di mercato', ma più in generale come potere economico e potere *tout court*, interessando i diritti fondamentali, i profili concorrenziali, il pluralismo e la stessa tenuta dei sistemi democratici. Si tratta pertanto di un fenomeno che merita attenzione da parte di tutte le istituzioni che contribuiscono a definire la *governance* dei mercati.

L'attuale assetto istituzionale è sostanzialmente adeguato a tutelare i diritti fondamentali, e in particolare il diritto alla protezione dei dati personali, e la concorrenza. Più complesso appare il tema

della protezione del pluralismo informativo nella moderna società digitale, in ragione di nuove dinamiche che, diversamente dagli approcci tradizionali al pluralismo, volti a disciplinare forme di accesso dal lato dell'offerta ai media tradizionali, sembrano riguardare, invece, i comportamenti degli

utenti dal lato della domanda, in un quadro di *overload* informativo e di limitata trasparenza circa l'origine delle informazioni e la loro natura editoriale, nonché circa gli effetti della profilazione sulla selezione dei contenuti proposti agli utenti.

Considerato il dinamismo e la complessità tecnica che caratterizzano gli ambiti presi in considerazione, Governo e Parlamento hanno la responsabilità di assicurare lo sviluppo equilibrato della cd. Economia digitale nel rispetto dei diritti e delle libertà fondamentali, nonché di interrogarsi sulla necessità di promuovere un appropriato quadro normativo che affronti la questione della piena ed effettiva trasparenza e liceità nell'uso dei dati personali.

2. Rafforzare la cooperazione internazionale sul disegno di policy per il governo dei Big Data.

La crescente interdipendenza dei mercati e di sistemi economici fa sì che le questioni sollevate dall'economia dei dati assumano spesso carattere sovra-nazionale.

Pertanto, in questo scenario nuovo ed evolutivo, un coordinamento fra le autorità della concorrenza europee non è solo auspicabile, ma necessario. A livello europeo, l'AGCM ha aderito alla Rete Europea della Concorrenza (*European Competition Network* - ECN), che riunisce la Commissione europea e le autorità *antitrust* istituite in ogni Stato Membro dell'Unione Europea, competenti ad applicare le regole di concorrenza stabilite dal TFUE. In particolare, nell'ambito dell'ECN, è stato costituito un gruppo di lavoro, denominato "*ECN Digital Markets",* ove le Autorità europee

espongono le attività in corso in merito all'applicazione delle regole di concorrenza relative ad operatori digitali. La costituzione di tale gruppo di lavoro è volta, da un lato, a promuovere la cooperazione tra le autorità degli Stati Membri, d'altro lato, a favorire la corretta allocazione di procedimenti istruttori riguardanti l'economia digitale. Inoltre, proprio anche in ragione della rapida evoluzione dei mercati digitali, è stata emanata la Direttiva UE 2019/1 (anche denominata ECN+) che conferisce alle autorità garanti della concorrenza degli Stati Membri poteri di applicazione più efficaci. Il dialogo transfrontaliero europeo non è soltanto legato a temi di concorrenza ma vi è anche un coordinamento interdisciplinare.

Infatti, l'AGCM partecipa alla *Digital Clearing House*, istituita su iniziativa dello *European Data Protection Supervisor* (EDPS) per valutare le implicazioni dei *Big Data* sotto il profilo della tutela del consumatore, della concorrenza e della protezione dei dati personali. In ambito extra europeo, la cooperazione multilaterale tra autorità Antitrust viene attuata in tre sedi principali: l'Organismo per la Cooperazione e lo Sviluppo Economico (OCSE), l'*International Competition Network* (ICN) e il *United Nations Conference on Trade and Development* (UNCTAD).

Sotto un diverso angolo visuale, nella misura in cui un trattamento di dati personali posto in essere mediante tecniche di *Big Data* ha natura transfrontaliera, trova invece piena applicazione la disciplina di protezione dei dati personali secondo il modello di cooperazione rafforzata tra autorità nazionali di protezione dei dati personali previsto dal Capo VII del RGPD (artt. 60 ss). Modalità ulteriori (ancorché meno stringenti) di cooperazione tra autorità di protezione dei dati possono altresì avere luogo, su scala globale, nell'ambito del *Global Privacy Enforcement Network* – GPEN.

Sotto il profilo regolamentare, l'Autorità per le Garanzie nelle Comunicazioni partecipa attivamente all'analisi del BEREC sulla Economia dei dati (*Data Economy*), finalizzata a conoscere l'impatto della economia dei dati sui mercati delle comunicazioni elettroniche, nonché quale ruolo possano avere i servizi di comunicazione elettronica per lo sviluppo della economia dei dati, ed, infine, in quale misura le autorità nazionali di regolazione (ANR) possano trarre vantaggio dall'economia dei dati nello svolgimento delle attività istituzionali. In questo contesto, il BEREC si prefigge di stimolare la collaborazione tra le ANR al fine di valutare in qual modo l'esperienza regolamentare fin qui acquisita possa rivelarsi utile per affrontare le possibili problematiche competitive connesse allo sviluppo della economia dei dati. In particolare, assume rilevanza il ruolo del BEREC nel tracciare le linee guida per le ANR sui temi della neutralità della rete che vanno aggiornati al fine di tener conto, oltre al ruolo dei prezzi espliciti (nelle pratiche di cosiddetto *zero rating*), anche dello scambio implicito di servizi a fronte del rilascio di permessi circa l'uso del dato personale laddove tale pratica venisse associata a forme di discriminazione nell'erogazione di medesimi servizi agli utenti. Appare inoltre necessario, sempre in ambito BEREC, affrontare il tema della proprietà dei dati generati nel contesto delle connessioni 5G e oggetto di scambio tra imprese attive in settori oggetto di regolazione distinta (energia, comunicazioni elettroniche, trasporti, sanità ecc.), nonché il tema associato della standardizzazione del dato e al fine di favorire l'interoperabilità dei servizi offerti.

Infine, anche AGCom partecipa alla *Digital Clearing House* dell'Unione Europea, in cui si esaminano le problematiche dei *Big Data*, con riguardo ad aspetti di interesse per tutte e tre le Autorità.

3. Promuovere una policy unica e trasparente circa l'estrazione, l'accessibilità e l'utilizzo dei dati pubblici al fine della determinazione di politiche pubbliche a vantaggio di imprese e

cittadini. Sarà necessario un coordinamento tra tale policy e le strategie europee già esistenti per la costituzione di un mercato unico digitale.

Il ricorso ai *Big Data* in modo crescente interessa trattamenti ulteriori rispetto a quelli effettuati mediante reti di comunicazioni elettronica ovvero nel settore privato, estendendosi ai trattamenti da parte di soggetti pubblici nel perseguimento di finalità istituzionali.

Anche tali soggetti, i quali peraltro originariamente acquisiscono le informazioni personali finalizzate all'assolvimento della propria missione istituzionale sulla base di obblighi legali gravanti sugli interessati, devono assicurarsi che il ricorso alle tecniche *Big Data*, anche con l'ausilio dei responsabili della protezione dei dati (Rpd), avvenga nel rispetto delle discipline di protezione dei dati personali.

4. Ridurre le asimmetrie informative tra utenti e operatori digitali, nella fase di raccolta dei dati, nonché tra le grandi piattaforme digitali e gli altri operatori che di tali piattaforme si avvalgono.

La riduzione dell'asimmetria informativa tra utenti e operatori digitali nella fase di raccolta dei dati costituisce un fondamentale obiettivo di *policy* al quale possono e devono contribuire diversi strumenti. In questo quadro appare rilevante informare compiutamente l'utente-consumatore non solo circa gli usi dei dati ceduti, ma anche circa la necessità della cessione in merito al funzionamento del servizio offerto. Il rapporto *interim* dell'AGCom ha evidenziato come molte *app* mostrino una relazione inversa tra prezzo di acquisto dell'*app* e permessi richiesti all'utente, talvolta anche per la medesima *app*. Appare indispensabile che l'utente, nelle decisioni di acquisto del servizio e di cessione del dato abbia piena consapevolezza della relazione tra permessi necessari al funzionamento dell'*app* e permessi ulteriori richiesti a seguito di cessione del dato.

Sia l'applicazione della normativa sulla protezione dei dati personali che la strumentazione propria della tutela del consumatore possono offrire un contributo importante per la riduzione di tale asimmetria informativa, garantendo che gli utenti ricevano un'adeguata, puntuale e immediata informazione circa le finalità della raccolta e dell'utilizzo dei loro dati e siano posti nella condizione di esercitare consapevolmente ed effettivamente le proprie scelte di consumo. In questa prospettiva, appaiono opportune misure volte a rendere maggiormente consapevoli i consumatori nel momento in cui forniscono il consenso al trattamento dei loro dati personali.

Appare altresì ineludibile che si proceda ad una progressiva riduzione delle asimmetrie informative tra le grandi piattaforme digitali e gli altri operatori che si avvalgono di tali piattaforme, aumentando la trasparenza dei criteri con i quali i dati vengono analizzati ed elaborati (ad esempio, nella definizione del *ranking* relativo al posizionamento e alla visibilità sulla piattaforma) e favorendo l'ingresso di nuovi intermediari dei dati che, su mandato degli utenti e nel rispetto della normativa a tutela della *privacy*, possano interfacciarsi con le grandi piattaforme globali con un accresciuto potere negoziale in merito alla contrattazione sul valore del dato e sul suo impiego commerciale.

In ogni caso, si avverte la necessità (non solo nello scenario nazionale) che le autorità di controllo siano messe in condizione di dotarsi di adeguati profili professionali (i cd. *data scientist*) per garantire l'adempimento dei propri compiti istituzionali.

5. Prima delle operazioni di trattamento dei dati, identificare la loro natura e proprietà e valutare la possibilità d'identificazione della persona a partire da dati 'anonimizzati'.

Appare necessario che chi intenda effettuare operazioni di trattamento secondo la metodologia propria dei *Big Data* si accerti, in via preliminare, della natura personale o meno dei dati trattati, così da identificare la cornice normativa di riferimento all'interno della quale opera.

Inoltre, i titolari del trattamento che intendono far uso di *Big Data* dovrebbero preventivamente valutare se una persona possa essere ragionevolmente identificata a partire dalla serie di dati "anonimizzata" utilizzata nel corso dell'analisi, in ragione delle operazioni di trattamento effettuate e dei *dataset* impiegati. Ciò non solo nell'ottica di rafforzamento della sicurezza del trattamento dei dati personali, come già previsto dal RGPD, ma anche nell'ottica di coerenza con la strategia nazionale di sicurezza cibernetica.

6. Introdurre nuovi strumenti per la promozione del pluralismo on-line, la trasparenza nella selezione dei contenuti nonché la consapevolezza degli utenti circa i contenuti e le informazioni ricevute on-line.

La concorrenza è senza dubbio uno strumento utile, ma insufficiente, per garantire la tutela del pluralismo. Anche un processo competitivo funzionante può, infatti, portare ad assetti di mercato incoerenti con un'informazione effettivamente pluralistica, in presenza di strategie di disinformazione nonché di fenomeni di autoselezione informativa, particolarmente diffusi nei comportamenti *on-line*, come il pregiudizio di conferma (*confirmation bias*), l'ancoraggio alle prime impressioni (*anchoring effect*), le camere d'eco (*echo chamber*), il conformismo di gruppo (*groupthink effect*) e così via. In questi casi, la disponibilità di una pluralità di fonti informative, e quindi l'operare del meccanismo concorrenziale dal lato dell'offerta, potrebbe non essere sufficiente a generare informazione verificata di qualità, diversità e pluralismo, potendo produrre anzi, in taluni casi, forme accentuate di autoselezione e polarizzazione nella ricerca e nella diffusione di informazioni (*backfire effect*).

Negli ultimi anni sia la Commissione europea che l'AGCom hanno avviato un percorso di autoregolazione e di co-regolamentazione che coinvolge tutte le componenti della società, ed in particolare tende a responsabilizzare le piattaforme tecnologiche, mediante l'adozione di appositi codici di comportamento volti a garantire sforzi concreti in favore della correttezza, completezza, verificabilità e non discriminatorietà dell'informazione accessibile *on-line*. Si tratta di un approccio distinto da quello tradizionalmente rivolto agli operatori media tradizionali in ragione della diversa natura di costruzione e diffusione dei contenuti, nonché della natura e della responsabilità editoriale degli stessi. Sul piano europeo, si ricorda il Piano d'adozione adottato a dicembre 2018 dalla Commissione europea per rafforzare la cooperazione tra Stati membri ed istituzioni europee, al fine di contrastare la crescente disinformazione che caratterizza il web e, tra le altre cose, condiziona negativamente la formazione del libero pensiero e le scelte del cittadino, in particolare, con riguardo alla formazione dell'orientamento politico ed all'espressione del voto.

Sul fronte delle iniziative dell'AGCom, si richiama in particolare il "*Tavolo per la garanzia del pluralismo e della correttezza dell'informazione sulle piattaforme digitali*" istituito nel novembre 2017, che ha l'obiettivo di favorire e promuovere l'autoregolamentazione delle piattaforme e lo scambio di buone prassi per l'individuazione ed il contrasto dei fenomeni di disinformazione *on-line*

frutto di strategie mirate. L'iniziativa, peraltro, si inscrive nel percorso istituzionale intrapreso da AGCom già a partire dal 2015, con la pubblicazione di rapporti e indagini conoscitive sul sistema dell'informazione *on-line*.

Dall'esperienza AGCom emergono tuttavia alcuni evidenti limiti di forme di autoregolazione che non siano accompagnate da poteri di *audit* e di *inspection* circa il ruolo della profilazione algoritmica nella selezione dei contenuti. Risulta pertanto auspicabile una verifica terza e indipendente degli esiti e dell'impatto misurabile delle iniziative di autoregolazione. In questa prospettiva, sembrano opportune iniziative legislative volte ad assicurare alle autorità indipendenti preposte alla tutela del pluralismo, poteri di *audit* e di *inspection* circa la profilazione algoritmica ai fini della selezione delle informazioni e dei contenuti, nonché in relazione agli esiti dell'applicazione delle *policy* e delle regole che le piattaforme digitali globali si sono date in tema di rimozione di informazioni false o di *hatespeech*.

Manca, infatti, ad oggi, una reportistica verificabile di tali autonome iniziative. Con riferimento, infine, al tema delle espressioni d'odio (*hatespeech*), a seguito della trasposizione della nuova Direttiva sui servizi media audiovisivi, l'AGCom applicherà anche alle piattaforme di *videosharing* il proprio regolamento di cui alla Delibera 157/19/CONS, avviando nel frattempo forme di co-regolazione per questo tipo di piattaforme.

7. Perseguire l'obiettivo di tutela del benessere del consumatore con l'ausilio degli strumenti propri del diritto antitrust estendendoli anche alla valutazione di obiettivi relativi alla qualità dei servizi, all'innovazione e all'equità.

Sotto il profilo dell'*enforcement* antitrust, la repressione di comportamenti abusivi da parte dei grandi *player* dell'economia digitale e di intese restrittive della concorrenza, entrambi facilitati dallo sviluppo di nuovi *software* e algoritmi sofisticati, è una delle priorità nell'attività dell'AGCM.

Le caratteristiche dell'economia digitale richiedono la ricerca di un nuovo equilibrio tra il rischio di scoraggiare i processi innovativi e il rischio di *under-enforcement*.

La capacità di profilazione, portata ai suoi estremi, e l'esasperazione degli effetti di rete possono agevolare comportamenti abusivi idonei a ridurre la contendibilità degli ecosistemi delle principali piattaforme, rendendo persistente il loro potere di mercato. In particolare, in ragione della natura multisettoriale dell'economia digitale e della presenza di grandi operatori digitali attivi su più mercati, la definizione del mercato rilevante ai fini dell'accertamento del potere di mercato potrebbe essere ripensata, talvolta tenendo significativamente in considerazione anche altri elementi.

In futuro la diffusione di algoritmi di prezzo pro-collusivi può facilitare la stabilità di cartelli e la creazione di contesti di mercato favorevoli ad equilibri collusivi.

L'AGCM intende prestare una particolare attenzione alle condotte delle piattaforme digitali che possono potenzialmente determinare effetti restrittivi della concorrenza, come dimostrano le istruttorie antitrust recentemente avviate.

In questa prospettiva, inoltre, per perseguire l'obiettivo di tutela del benessere del consumatore, diventa opportuno non confinare l'analisi ai tradizionali parametri legati a prezzi e quantità, ma, con l'ausilio degli strumenti propri del diritto *antitrust*, estenderla anche alla qualità, all'innovazione e all'equità.

Appaiono infine necessarie, quantomeno con riferimento alle piattaforme digitali globali, misure volte ad incrementare la trasparenza all'utente circa la natura della propria profilazione in merito ai contenuti ricevuti, nonché meccanismi di o*pt-in* circa il grado di profilazione prescelto, e ciò anche ai fini della tutela del pluralismo *on-line*, in relazione alla selezione dei contenuti operante attraverso la profilazione del consumatore.

8. Riformare il controllo delle operazioni di concentrazioni al fine di aumentare l'efficacia dell'intervento delle autorità di concorrenza.

Con la diffusione dei *Big Data*, il controllo delle concentrazioni assume una nuova centralità. Al fine di aumentare l'efficacia dell'intervento delle autorità di concorrenza rispetto alle operazioni di concentrazione è auspicabile:

1. una riforma a livello nazionale e internazionale che consenta alle autorità di concorrenza di poter valutare pienamente anche quelle operazioni di concentrazione sotto le attuali soglie richieste per la comunicazione preventiva, ma che potrebbero risultare idonee a restringere sin dalla loro nascita importanti forme di concorrenza potenziale (come le acquisizioni da parte dei grandi operatori digitali di *start-up* particolarmente innovative anche soprannominate '*killing acquisitions*');

2. la modifica dell'art. 6, comma 1, della legge n. 287/90, con l'introduzione di uno *standard* valutativo più adatto alle sfide dell'economia digitale, che faccia leva sul criterio dell'impedimento significativo della concorrenza effettiva (SIEC – "*Substantial impediment to effective competition*").

9. Agevolare la portabilità e la mobilità di dati tra diverse piattaforme, tramite l'adozione di standard aperti e interoperabili

Agevolare la portabilità e la mobilità di dati tra diverse piattaforme, tramite l'adozione di standard aperti e interoperabili, anche oltre quanto già previsto dal diritto alla portabilità di cui all'art. 20 del RGPD, costituisce un obiettivo con una forte valenza pro-concorrenziale.

In casi particolari, ferma restando la necessità di tutelare il diritto alla protezione dei dati personali, la tutela della concorrenza potrebbe richiedere obblighi di mobilità e portabilità dei dati personali ulteriori rispetto a quelli previsti in generale dal RGPD.

A questo riguardo, si dovrebbe considerare la possibilità di estendere lo strumento della portabilità dei dati, oltre quanto – meritoriamente – stabilito dall'articolo 20 del RGPD, prevedendo una disciplina della portabilità dei dati, che favorisca lo sviluppo della competizione nei vari ambiti di valorizzazione economica del dato e, di conseguenza, una più efficace tutela del consumatore-utente. Possono pertanto essere prese in considerazione iniziative legislative o regolamentari, nell'ambito della cooperazione con l'Unione Europea, per disciplinare l'interoperabilità delle piattaforme tecnologiche, così da consentire effettivamente all'utente una piena portabilità dei propri dati.

10. Rafforzare i poteri di acquisizione delle informazioni da parte di AGCM ed AGCom al di fuori dei procedimenti istruttori e aumento del massimo edittale per le sanzioni al fine di garantire un efficace effetto deterrente delle norme a tutela del consumatore.

La tutela del consumatore può intervenire su una molteplicità di profili connessi al rapporto tra operatori e utenti nella fase di acquisizione, elaborazione e trattamento dei dati. La circostanza che

alle condotte poste in essere dalle imprese sia applicabile la normativa in materia di protezione dei dati personali non le esonera dal rispettare le norme in materia di pratiche commerciali scorrette, ponendosi le due discipline su un piano di complementarietà e non di alternatività. Le caratteristiche

delle politiche di protezione dei consumatori e di tutela della *privacy* sono senza dubbio componenti importanti di un confronto concorrenziale *fair*.

Tenuto conto delle grandi dimensioni di molti operatori attivi nell'economia digitale, impregiudicato quanto già previsto dal RGPD, al fine di garantire un efficace effetto deterrente delle norme a tutela del consumatore, sembra necessario prevedere quanto prima un aumento del massimo edittale per le sanzioni.

Al fine di consentire una piena comprensione dei nuovi fenomeni in atto nell'economia digitale, appare opportuno il rafforzamento dei poteri di acquisizione delle informazioni da parte di AGCM ed AGCom al di fuori dei procedimenti istruttori (indagini conoscitive, attività pre-istruttoria), anche prevedendo la possibilità di irrogare sanzioni amministrative in caso di rifiuto o ritardo nel fornire le informazioni richieste o in presenza di informazioni ingannevoli od omissive. In questa direzione è peraltro orientata la cornice normativa in materia di protezione dei dati personali (cfr. artt. 157 e 166, comma 2, del Codice in materia di protezione dei dati personali).

11. Istituzione di un "coordinamento permanente" tra le tre Autorità

Un'efficace politica pubblica per i *Big Data* e l'economia digitale richiede non solo l'*enforcement*, ma anche un'adeguata attività di *advocacy* di cui l'iniziativa congiunta tra AGCM, AGCom e Garante per la protezione dei dati personali è testimonianza – volta a:

- contrastare le norme e le regolazioni volte a proteggere assetti di mercato "maturi" a scapito dello sviluppo delle innovazioni favorite dalla digitalizzazione, che contribuiscono alla competitività del sistema economico e al benessere dei consumatori;
- definire un *level playing field* attraverso misure volte alla rimozione degli ingiustificati vantaggi sotto il profilo fiscale e delle relazioni industriali di cui beneficiano i principali protagonisti della rivoluzione digitale in generale e in relazione ai diversi mercati rilevanti interessati e ai versanti intermediati dalle grandi piattaforme digitali;
- emancipare e accrescere la consapevolezza della collettività sia dei benefici che dei rischi derivanti dalla digitalizzazione dell'economia.

Con riguardo al tema specifico dell'accesso ai dati, le sinergie tra tutela della concorrenza e regolazione possono risultare preziose:

- ai sensi della normativa *antitrust*, un'impresa in posizione dominante può essere soggetta all'obbligo di fornire accesso ai dati indispensabili e non agevolmente duplicabili per salvaguardare la concorrenza in uno o più mercati in cui la medesima impresa è attiva;
- se invece l'obiettivo sociale è quello di tutelare interessi pubblici diversi dalla promozione della concorrenza, eventuali circoscritti interventi regolatori in materia di accesso ai dati appaiono più efficaci, così come possono contribuire alla promozione della concorrenza quando l'intervento *antitrust* si riveli insufficiente;
- in particolare, con riferimento alle competenze AGCom, sarà oggetto di valutazione quanto previsto dal nuovo codice europeo delle comunicazioni elettroniche di cui alla Direttiva UE

2018/1972 dell'11 dicembre 2018, laddove si specifica che "il trattamento dei dati personali da parte dei servizi di comunicazione elettronica, sia esso in forma di remunerazione o in altra forma, dovrebbe essere conforme al regolamento (UE) 2016/679 del Parlamento europeo e del Consiglio" (par. 15) e che "i servizi di comunicazione elettronica sono spesso forniti all'utente finale non solo in cambio di denaro, ma in misura sempre maggiore e in particolare in cambio della comunicazione di dati personali o di altri dati" concludendo che il concetto di remunerazione dovrebbe pertanto ricomprendere anche "le situazioni in cui l'utente finale è esposto a messaggi pubblicitari come condizione per l'accesso al servizio o le situazioni in cui il fornitore del servizio monetizza i dati personali raccolti in conformità del regolamento (UE) 2016/679" (par. 16);

- laddove l'accesso ai dati vada garantito nell'interesse generale, regolazioni settoriali che consentano allo Stato di accedere a banche dati raccolte da imprese private e utili per ragioni di salute pubblica, ambientali, di sicurezza o di mobilità, sembrano lo strumento più appropriato per garantire obiettivi di interesse pubblico ed evitare inutili e costose duplicazioni di dati già disponibili;
- l'accesso a taluni dati in possesso delle piattaforme digitali globali, e la loro replicabilità, può essere necessario anche per i soggetti preposti alla rilevazione delle 'audience', al fine di garantire lo sviluppo pro-concorrenziale dei mercati della pubblicità *on-line* - e del cosiddetto *programmatic advertising* basato sulla profilazione algoritmica degli utenti - e di assicurare un'equa ripartizione delle risorse idonea a promuovere un'offerta informativa di qualità;
- in ogni caso, eventuali interventi regolatori in materia di accesso ai dati devono essere necessari e proporzionati e devono tenere in considerazione le specificità del servizio/mercato al quale si riferiscono nonché le finalità sociali ad essi connesse e oggetto di presidio regolatorio;
- il contenuto di eventuali obblighi di accesso ai dati personali – in termini di ampiezza, natura e modalità – deve essere adeguatamente bilanciato con il diritto alla protezione dei dati personali.

Più in generale, le sfide poste dallo sviluppo dell'economia digitale e dai *Big Data* richiedono uno sfruttamento pieno delle sinergie esistenti tra strumentazione *ex ante* ed *ex post*, a tutela della *privacy*, della concorrenza, del consumatore e del pluralismo.

AGCM, AGCom e Garante per la protezione dei dati personali, ciascuno nell'ambito delle proprie competenze, possono meglio garantire i propri obiettivi istituzionali, nella misura in cui sapranno cogliere a pieno le opportunità offerte da una proficua cooperazione.

A tal fine, le tre Autorità, nell'esercizio delle competenze complementari ad esse assegnate e che contribuiscono a fronteggiare le criticità dell'economia digitale, si impegnano a strette forme di collaborazione negli interventi che interessano i mercati digitali, anche attraverso la sottoscrizione di un *memorandum of understanding*.

Audizioni con soggetti che operano nei diversi settori interessati dallo sviluppo dei *Big Data*

- Allianz SE, in data 17 novembre 2017;

- GEDI Gruppo Editoriale S.p.A., in data 21 novembre 2017;

- RAI-Radio Televisione Italiana S.p.A., in data 24 novembre 2017;

- Gruppo Mediaset S.p.A., in data 28 novembre 2017;

- Experian Italia S.p.A., in data 28 novembre 2017;

- IlSole24ore S.p.A., in data 1dicembre 2017;

- CRIF S.p.A., in data 18 dicembre 2017;

- Facebook Italy S.p.A., in data 5 febbraio 2018;

- Intesa San Paolo S.p.A., in data 23 febbraio 2018;

- Unicredit S.p.A., in data 8 marzo 2018;

- Generali S.p.A., in data 21 marzo 2018;

- Microsoft Corporation, in data 9 ottobre 2018;

- IBM Italia S.p.A., in data 22 ottobre 2018;

- Amazon Italia Services S.r.l., in data 26 novembre 2018;

- Wind-Tre S.p.A., in data 29 novembre 2018;

-Tim S.p.A., in data 7 dicembre 2018;

-Vodafone Italia S.p.A., in data 7 dicembre 2018;

- Fastweb S.p.A., in data 7 dicembre 2018.

Audizioni con professori universitari ed esperti del settore:

- Dott. C. Giustozzi dell'AGID, in data 16 novembre 2017;

- Prof. A. Mantelero del Politecnico di Torino, Prof.ssa A. Papa dell'Università degli Studi di Napoli Parthenope e Prof.ssa V. Falce dell'Università Europea di Roma in data 21 novembre 2017;

- Prof.ssa F. Giannotti e Prof. D. Pedreschi dell'Università di Pisa - KDD Lab, in data 5 dicembre 2017;

- Prof. M.Gambaro dell'Università degli Studi di Milano, in data 18 dicembre 2017;

- Prof. J. Cannataci dell'University of Malta, in data 15 gennaio 2018;

- Prof.ssa T. Maggiolino dell'Università Commerciale Luigi Bocconi ed il Prof. A. Preta dell'IT Media Consulting, in data 30 gennaio 2018;

- Prof. A. De Streel dell'Université de Namur, in data 19 febbraio 2018;

- Dott. S. Quintarelli, in data 13 settembre 2018.