

IL RAPPORTO TRA BIG DATA E DIRITTO ALLA TUTELA DEI DATI PERSONALI

AUTORI

Aurora Agostini, Counsel

Jessica Giussani, Associate

Data & Technology Innovation Team

SOMMARIO

Il rapporto tra Big Data e diritto alla tutela dei dati personali	2
Quali sono le caratteristiche dei big data?.....	2
Il processo di estrazione di conoscenza.....	1
1. La raccolta dei Big Data	1
2. L'elaborazione dei Big Data.....	1
3. L'interpretazione e l'utilizzo dei Big Data.....	1
L'uso dei big data nel quotidiano.....	2
È possibile tutelare la privacy dell'utente?	2
Il valore dei big data per il business.....	2
GDPR e Big Data	3
Gli interessati sono proprietari dei loro dati?.....	4
L'anonimizzazione può essere la soluzione?.....	4
Le novità del Digital Service Package	5

IL RAPPORTO TRA BIG DATA E DIRITTO ALLA TUTELA DEI DATI PERSONALI

In tempi recenti, la raccolta e l'analisi dei cosiddetti Big Data ha profondamente cambiato il modo di fare impresa, rendendo il loro impiego quasi imprescindibile per le società che intendono garantirsi un vantaggio competitivo. Infatti, i Big Data forniscono alle società importanti informazioni sulle abitudini di consumo dei propri clienti, che possono essere utilizzate da quest'ultime per orientare il proprio business. Lo sviluppo di un'economia sempre più digitale, fondata sulla raccolta e analisi di un'ingente mole di dati in continua espansione, rende i Big Data uno dei settori più rilevanti e discussi del diritto di Internet.

Sebbene il concetto di Big Data sia relativamente nuovo, le origini dei grandi sistemi di dati risalgono alla metà del secolo scorso: già nel 1950 un Cinque anni prima, durante la seconda guerra mondiale, grazie alle teorie di Turing era stato sviluppato uno dei primi computer (il Colossus), con lo scopo di individuare e decifrare i messaggi intercettati a una velocità impensabile per l'epoca, riducendo il tempo necessario a svolgere questa attività da settimane a poche ore.

Negli anni '60, grazie allo sviluppo dei primi data center, fu per la prima volta possibile digitalizzare e archiviare grandi quantità di dati, anche se in server molto lontani per dimensione e potenza rispetto a quelli attuali.

articolo del ricercatore Alan Turing si apriva con la provocazione molto forte: "Can machines think".

Infine, negli anni '90, con la nascita del web si posero le basi per la vera e propria rivoluzione dei Big Data. Solo intorno al 2005, con l'avvento dei primi social network, si è compreso però il reale potenziale dell'ingente quantità di dati generati dagli utenti attraverso Facebook, YouTube e altri servizi online:

nello stesso anno è stato sviluppato il primo software creato appositamente per archiviare e analizzare grandi insiemi di dati - Apache Hadoop - la cui versione aggiornata è ancora oggi utilizzata per l'elaborazione dei Big Data da società come Ebay, Facebook, LinkedIn, Twitter e Spotify.

Quali sono le caratteristiche dei big data?

Con Big Data si intende un insieme di dati di rilevanti dimensioni provenienti da diverse fonti, tra le quali, smartphone, social media, wearable ecc. Tali dati possono presentarsi sia in forma strutturata, più facilmente analizzabile, che non strutturata o semi-strutturata (per cui si rendono necessari sistemi di elaborazione del dato più complessi).

I Big Data possono inoltre essere definiti attraverso le cd. tre "V":

- ▶ volume dei dati generati e raccolti;
- ▶ varietà delle tipologie di dati disponibili (tanto più variano le fonti produttive, tanto più variano i dati disponibili);
- ▶ velocità delle operazioni di trattamento.

Oltre a queste, sono state individuate ulteriori V, tra cui il valore che i dati hanno una volta elaborati ed analizzati.

Il processo di estrazione di conoscenza

La finalità ultima degli articolati processi sottesi all'utilizzo di Big Data è quella di migliorare l'efficienza dei processi produttivi e orientare le scelte strategiche di chi gestisce un business. Ciò avviene grazie ad una individuazione più accurata delle tendenze di mercato, e quindi del target di consumatore ideale per un determinato prodotto o servizio, nonché una profilazione più mirata della pubblicità e delle proposte commerciali. A tal fine è cruciale il processo di "estrazione di conoscenza" dai Big Data, che si sviluppa in 3 fasi:

1. la raccolta;
2. l'elaborazione,
3. l'interpretazione e utilizzo.

1. LA RACCOLTA DEI BIG DATA

Le fonti di provenienza dei dati sono molteplici. Soprattutto nel contesto attuale, in cui tutti i contenuti sono resi disponibili in formato digitale e la maggior parte delle attività viene svolta online, è facile reperire dagli utenti grandi quantità di dati. Si pensi, ad esempio, ai dispositivi personali come smartphone, tablet e computer, alla navigazione satellitare, ai social network, in cui i fruitori pubblicano i propri contenuti (foto, video, testi), alle app e ai siti internet.

A ciò si aggiunge l'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 indossabili (ad esempio, i wearable device che collezionano dati relativi alle attività sportive e/o ai parametri biologici).

2. L'ELABORAZIONE DEI BIG DATA

I dati isolati hanno poco valore, ma lo acquisiscono quando sono organizzati: per tale ragione, la fase dell'elaborazione riveste un ruolo centrale nell'intera filiera dei Big Data, perché permette l'organizzazione dei dati grezzi non strutturati in informazioni suscettibili di essere utilizzate per finalità economiche.

L'informazione è il risultato del processo di analisi del dato.

3. L'INTERPRETAZIONE E L'UTILIZZO DEI BIG DATA

Per la prima volta nella storia dell'umanità siamo in grado di tenere traccia di quello che milioni di persone fanno ogni giorno. Riportiamo di seguito qualche esempio pratico di quanto i Big Data siano un fenomeno che interessa da vicino la nostra vita quotidiana:

- ▶ Gli smartwatch che, attraverso l'analisi dei passi effettuati e del battito cardiaco, tengono traccia delle varie attività svolte durante la giornata. Questi strumenti sono in grado, ad esempio, di individuare quando l'individuo che li indossa è sotto stress.
- ▶ Spotify non si limita a tenere traccia dei gusti musicali degli utenti, infatti, attraverso l'analisi delle playlist che l'utente riproduce durante la giornata, l'App è in grado di riconoscere quando è felice, quando è in viaggio o quando sta facendo una festa.
- ▶ le App di delivery che raccolgono e vendono dati, relativi alle abitudini di consumo degli utenti, ai ristoranti nella zona di interesse dell'utente. I ristoranti potranno quindi utilizzare tali informazioni per la strutturazione di attività commerciali e di marketing come, ad esempio, l'invio di codici sconto o promozioni.

- ▶ Le piattaforme di car/bike/scooter sharing che, grazie alla geolocalizzazione, identificano i luoghi frequentati dagli utenti.
- ▶ Gli strumenti di pagamento elettronici che consentono di acquisire informazioni sui comportamenti di acquisto e le preferenze degli utenti che li utilizzano, utili a verificare l'efficacia di campagne pubblicitarie personalizzate, nonché a profilare ulteriormente i propri utenti.

L' utilizzo dei big data nel quotidiano

Per la prima volta nella storia dell'umanità siamo in grado di tenere traccia di quello che milioni di persone fanno ogni giorno. Riportiamo di seguito qualche esempio pratico di quanto i Big Data siano un fenomeno che interessa da vicino la nostra vita quotidiana:

- ▶ Gli smartwatch che, attraverso l'analisi dei passi effettuati e del battito cardiaco, tengono traccia delle varie attività svolte durante la giornata. Questi strumenti sono in grado, ad esempio, di individuare quando l'individuo che li indossa è sotto stress.
- ▶ Spotify non si limita a tenere traccia dei gusti musicali degli utenti, infatti, attraverso l'analisi delle playlist che l'utente riproduce durante la giornata, l'App è in grado di riconoscere quando è felice, quando è in viaggio o quando sta facendo una festa.
- ▶ Le App di delivery che raccolgono e vendono dati, relativi alle abitudini di consumo degli utenti, ai ristoranti nella zona di interesse dell'utente. I ristoranti potranno quindi utilizzare tali informazioni per la strutturazione di attività commerciali e di marketing come, ad esempio, l'invio di codici sconto o promozioni.
- ▶ Le piattaforme di car/bike/scooter sharing che, grazie alla geolocalizzazione, identificano i luoghi frequentati dagli utenti.
- ▶ Gli strumenti di pagamento elettronici che consentono di acquisire informazioni sui comportamenti di acquisto e le preferenze degli utenti che li utilizzano, utili a verificare l'efficacia di campagne pubblicitarie personalizzate, nonché a profilare ulteriormente i propri utenti.

È possibile tutelare la privacy dell'utente?

In definitiva, lo strumento che favorisce maggiormente il tracciamento dei dati degli utenti è lo smartphone. Esso riveste un ruolo centrale nell'acquisizione dei dati, in quanto è connesso a Internet, accompagna l'utente in tutte le sue attività quotidiane e dispone di numerosi dispositivi di input (come i sensori di movimento, di luminosità, di localizzazione, la tastiera e il touch screen).

Questa ingente quantità di dati creata dall'utilizzo degli smartphone è possibile grazie a due strumenti principali:

1. il sistema di geolocalizzazione che, anche se non attivato, traccia comunque la posizione dell'utente.
2. Le applicazioni che richiedono spesso l'accesso ai contatti, al microfono, alle fotografie e ad altre funzionalità in grado di tracciare l'utente, anche se non necessariamente pertinenti all'utilizzo dell'applicazione stessa, e che potranno essere oggetto di cessione a terzi.

Il valore dei big data per il business

Il valore dei Big Data può essere facilmente compreso facendo riferimento alla cosiddetta Data Economy: oggi infatti i dati costituiscono un inestimabile valore per coloro che riescono ad estrarli, utilizzarli e monetizzarli.

Esistono principalmente due business model per creare profitto dallo sfruttamento dei Big Data:

1. raccogliendo, elaborando ed interpretando i dati relativi ai propri utenti al fine di migliorare il proprio servizio e/o per specifiche finalità della società e quindi utilizzarli internamente.
2. Vendendo i dati a terzi. Ciò implica l'ingresso in scena della figura del Data Broker, ovvero professionisti (spesso società) che si occupano di recuperare dati e informazioni per poi elaborarli, interpretarli, analizzarli al fine di creare un profilo. Tale profilo può essere venduto a società terze interessate a profilare gli utenti per meglio individuare chi rientra nel target della propria attività e personalizzare di conseguenza le campagne pubblicitarie. È bene sottolineare che i dati non vengono solo venduti a società che vendono prodotti di consumo ma anche a società finanziarie che li utilizzano per individuare i profili di rischio e i tassi di interesse da applicare ad un potenziale cliente.

GDPR e Big Data

Come l'impiego dei Big Data convive quindi con le norme in materia di protezione dei dati personali?

Il problema si pone sin dal momento della raccolta dei Big Data. Nell'enorme varietà e quantità di informazioni reperite può capitare che si raccolgano dati di natura personale. In tal caso il trattamento non può prescindere - almeno in Europa - dal rispetto del GDPR.

Tuttavia, è spesso problematico (i) da un lato, definire la linea di demarcazione tra dati di natura personale, in particolare in ragione della possibilità di riconnettere informazioni apparentemente anonime a individui singoli grazie, ad esempio, alle aumentate capacità di calcolo e alla pluralità di archivi in ipotesi utilizzabili. Si pensi ai dati pseudonimizzati - esempio indirizzo IP, nascosto solo in parte; (ii) dall'altro lato, l'acquisizione massiva dei dati rende infatti difficoltosa la specifica individuazione ex ante delle finalità del relativo trattamento.

Il GDPR prevede che le attività di raccolta e impiego dei dati possano avvenire previa richiesta del consenso dell'interessato o al ricorrere di una delle condizioni previste dall'art. 6. È inoltre stabilito 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). Tali dati devono essere anche 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. Infine, devono essere trattati in modo da garantirne un'adeguata sicurezza.

I principi di minimizzazione, limitazione della finalità e conservazione per il solo tempo indispensabile alla realizzazione del trattamento non si adattano però a raccolte massive. I dati vengono infatti raccolti non per esigenze attuali ma in vista di future ed eventuali necessità e riutilizzati per fini non sempre compatibili con quelli originari.

Sicuramente l'avvento del GDPR ha reso più trasparente il trasferimento di dati, tuttavia non è ancora sufficiente. Se è vero che l'utente deve fornire il proprio consenso e ha la facoltà di conoscere a chi questi dati vengano venduti, nella realtà dei fatti, pochissime persone hanno tempo o competenze per leggere e verificare le informative che gli vengono proposte ogni volta che accedono ad un sito. L'utente medio accetta quello che gli viene proposto per accedere nel minor tempo possibile al contenuto di suo interesse: si tratta del cosiddetto privacy paradox.

Un ulteriore problema riguarda il fatto che spesso l'utente non sembra avere alternative: sebbene sia informato o parzialmente informato, il rifiuto a fornire determinati dati può compromettere la fruizione dell'esperienza completa del sito, applicazione ecc.

In ogni caso, ricordiamo che se una società ha raccolto i dati utilizzando come base giuridica il consenso dell'interessato o un obbligo previsto dalla legge, non è ammesso alcun ulteriore trattamento al di fuori dei settori coperti dal consenso originale o dalla disposizione di legge.

Per questo motivo, i titolari del trattamento di dati personali, anziché raccogliere il consenso espresso dei propri utenti/clienti, spesso preferiscono ricorrere alla base giuridica del legittimo interesse, i cui limiti di demarcazione non sono ancora definiti in modo chiaro.

Per il momento, una soluzione possibile potrebbe essere quella del dynamic consent: 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. Si tratta però di un modello che per ora ha trovato applicazione solo per le bio-banche e per i consensi sulle ricerche bio-mediche.

Gli interessati sono proprietari dei loro dati?

Gli interessati hanno sempre il diritto di accedere, rettificare e cancellare i dati personali e di limitarne il trattamento ai sensi del GDPR: ciò significa che le società devono essere in grado di scavare nella grande quantità di dati archiviati in diversi sistemi per individuare e/o cancellare i dati appartenenti all'interessato. Ovviamente, l'interessato non ha la possibilità di verificare che il titolare abbia effettivamente eliminato tali dati, né è in grado di sapere se la cancellazione abbia riguardato anche quei dati che eventualmente siano stati ceduti a terzi.

Inoltre, molte società sostengono di operare in conformità al GDPR e di utilizzare i dati personali in modo lecito e corretto, fornendo all'interessato un'informativa puntuale e completa, ma non di rado dimenticano di cancellare i dati inutilizzati o il cui periodo di conservazione è terminato. Anche questa condotta costituisce una violazione del GDPR: l'articolo 5 stabilisce infatti che i dati personali non debbano essere conservati più a lungo del necessario. Le società non possono quindi limitarsi a stabilire periodi di conservazione, ma devono anche assicurarsi dell'effettiva cancellazione dei dati, a prescindere da una esplicita richiesta dell'interessato.

L'anonimizzazione può essere la soluzione?

Una possibile soluzione al problema dell'utilizzo dei dati personali contenuti nei Big Data (i.e. quando sono trasferiti a terzi oppure conservati più del tempo previsto) potrebbe essere quella di renderli anonimi. Sul punto, è bene tenere in considerazione che esiste una profonda differenza tra pseudonimizzazione e anonimizzazione.

La pseudonimizzazione consiste nella sostituzione degli identificatori diretti, in modo tale che i dati non possano più essere attribuiti a un soggetto specifico senza l'utilizzo di informazioni aggiuntive. Questo significa che un dato pseudonimizzato rimane comunque un dato personale; e come tale, soggetto alla disciplina del GDPR.

L'anonimizzazione, invece, si riferisce alla pratica di rendere i dati non identificabili in modo tale che sia impossibile ricostruire l'identità dell'interessato.

In linea teorica l'anonimizzazione potrebbe essere una soluzione nel caso in cui le società fossero certe dell'impossibilità di identificare nuovamente l'individuo. Nella realtà dei Big Data, ciò difficilmente è possibile; infatti, gli algoritmi utilizzati sono in grado di individuare, comparando varie banche dati, e re-identificare un utente. Dunque, anche l'anonimizzazione - fino a poco tempo fa ritenuta una tecnica affidabile - ormai nell'era dei Big Data non sembra essere più sufficiente a garantire l'effettiva riservatezza dei dati.

Le novità del Digital Service Package

Nel prossimo futuro lo scenario potrebbe cambiare. L'entrata in vigore del Digital Service Package, composto da due regolamenti, il Digital Services Act e Digital Markets Acts, introduce un regime speciale per i cosiddetti Gatekeepers con riferimento, tra le altre materie, alla protezione dei dati personali.

I Gatekeepers del mercato digitale sono i fornitori di servizi di piattaforme di base, quali social network, browser, motori di ricerca e servizi di messaggistica e possono venire individuati sulla base di tre diversi parametri:

- ▶ dimensione dell'impresa: fatturato annuo uguale o superiore a 7,5 miliardi di euro negli ultimi tre esercizi finanziari o valore totale delle azioni di mercato di almeno 7,5 miliardi nell'ultimo anno e fornitura di servizi di piattaforma ad almeno tre Stati dell'UE;
- ▶ controllo del gateway di accesso ai dati degli utenti: viene valutata la registrazione di almeno 10.000 utenti europei attivi durante l'ultimo anno e più di 45 milioni di utenti europei attivi al mese;
- ▶ posizione durevole e stabile sul mercato, se le soglie individuate nel precedente criterio sono state raggiunte in ciascuno degli ultimi tre esercizi finanziari.

Questi regolamenti sono fondamentali in tema Big Data, in quanto i Gatekeeper si identificano in quella manciata di società che gestisce i dati di quasi tutto il globo, tra le altre, Meta, Google, Apple, Amazon, Microsoft.

Alcune delle principali novità introdotte dai due regolamenti mirano a porre restrizioni sulla pubblicità basata sul monitoraggio dell'utenza, sul targeting e sulla profilazione. In particolare, il contenuto della normativa vieta ai Gatekeepers di:

- ▶ profilare i minori per scopi commerciali;
- ▶ processare i dati raccolti tramite terze parti che si servono della piattaforma del Gatekeeper per offrire servizi di pubblicità;
- ▶ combinare i dati personali raccolti sulla piattaforma con quelli raccolti su qualsiasi altra piattaforma del Gatekeeper o di terze parti o con quelli provenienti da altri servizi offerti separatamente dallo stesso Gatekeeper;
- ▶ iscrivere automaticamente l'utente ad altri servizi del Gatekeeper per combinare i dati personali.

In conclusione, i recenti sviluppi normativi ci fanno pensare che si sia finalmente acceso un faro sui protagonisti di questo mercato che fino ad ora avevano agito secondo il motto del "too big to care", troppo grande per curarsene. Il messaggio che le autorità competenti stanno cercando di trasmettere è che Big Data e Privacy possono e devono convivere, facendo però sempre rispettare la volontà esplicita di un consumatore adeguatamente formato nel comprendere l'importanza dei propri dati e del loro utilizzo.

Avv. Aurora
Agostini
[Contatti](#)

Dott.ssa Jessica
Giussani
[Contatti](#)

www.lexia.it

Via del Lauro 9,
20131 Milano (MI)

[Follow us on linkedin](#)