Big data, algoritmi e diritto

di Fernanda Faini

Abstract: Big data, algorithms and law – Data form our digital identities and are the foundation of every activity carried out online, constituting the basis on which the digital society stands. In the light of the European Union legal acts of reference, the contribution, after analyzing the big data, examines the emerging ethical-social questions and the legal issues involved, relating to ownership, responsibility, competition, rights protection and personal data protection. In consideration of the critical issues, the contribution outlines the principles and legal instruments suitable for achieving an ethical and legal governance of big data, firmly based on transparency, openness and collective control.

Keywords: Big data; Algorithms; Transparency; Ethics; Law.

1. Big data e algoritmi: il *framework* giuridico europeo

La società contemporanea poggia sulla pervasiva centralità dei dati quali risorse essenziali per lo sviluppo umano sotto i profili economico, sociale e culturale, "materia prima" di cui si nutre la tecnologia[1]. Le identità personali e le attività umane trovano fondamento in una moltitudine di dati, che, animati da potenti algoritmi, formano il nostro mondo digitale. La società odierna, infatti, tende a convertire i fenomeni e, più ampiamente, l'intera realtà in dati e a produrre *big data* analizzabili da algoritmi, capaci di generare valore proprio grazie all'elaborazione di enormi quantità di dati eterogenei[2].

In tale contesto di riferimento, pertanto, elementi privilegiati per osservare la nostra realtà sono costituiti dalle configurazioni odierne assunte dai dati, i *big data*, e dagli algoritmi, protagonisti indiscussi dell'evoluzione della conoscenza in senso dinamico e "attivo".

Al fine di esaminare le questioni che incontra il diritto nel momento in cui è chiamato ad affrontare *big data* e algoritmi, è necessario analizzare le caratteristiche specifiche di tali fenomeni e i mutamenti che inducono nel modo stesso di osservare la realtà da parte dell'uomo e, in specifico, del giurista.

I *big data* sono enormi volumi di dati detenuti da grandi organizzazioni come governi e multinazionali, provenienti da diverse fonti e analizzati per

1 M. Castells, *The rise of the Network society*, Oxford, 2000.

2 Cfr. V. Mayer-Schönberger, K. Cukier, *Big data. Una rivoluzione che trasformerà il nostro modo di vivere e che già minaccia la nostra libertà*, trad. it., Milano, 2013, 103 ss.; C. Accoto, *Il mondo dato. Cinque brevi riflessioni di filosofia digitale*, Milano, 2017, 1 ss.

mezzo di algoritmi e specifiche tecniche quali *data mining*, *big data analytics*, *machine learning*[3]. Si tratta di dati forniti volontariamente dagli utenti alle piattaforme digitali; dati "scambiati" a fronte di utilità (raccolte punti, sconti); dati forniti in modo più o meno consapevole (GPS, sensori); dati residui (*data exhaust*) o inferiti da altri dati; dati raccolti dai soggetti pubblici; i dati dell'*Internet of Things* (IoT), come quelli prodotti da dispositivi indossabili, "case intelligenti", automobili autonome[4]; i dati su cui si basano le soluzioni di intelligenza artificiale[5].

Gli algoritmi sono capaci di strutturare le informazioni e automatizzare i processi[6]: i *big data* sono caratterizzati dalla varietà, dal volume e dalla velocità, che sottende la rilevanza del tempo e della dinamicità[7]. Accanto a queste dimensioni, sono considerate caratteristiche dei *big data* anche il valore (quanto i *big data* valgono come insieme) e la veracità o veridicità (la qualità e l'accuratezza dell'analisi)[8].

Gli algoritmi si affidano alle correlazioni e alle inferenze e poggiano sulla probabilità, portando gli uomini che se ne servono a rinunciare alle ipotesi predeterminate e alla ricerca dei nessi di causalità. In altri termini, si presta attenzione alle correlazioni che emergono dalle analisi sui dati, senza che siano necessariamente predefiniti l'oggetto di indagine e gli obiettivi perseguiti[9].

L'interesse rivolto ai *big data* si collega strettamente alle caratteristiche che li connotano e a ciò che permettono di ottenere. La conoscenza che *big data* e algoritmi permettono di raggiungere sui fenomeni oggetto di osservazione, infatti, non riguarda solo il passato e il presente, ma anche il futuro, per mezzo di una vera e propria capacità predittiva: elevate correlazioni sono capaci di indicare

[3] Cfr. S. Faro, N. Lettieri, Big Data*: una lettura informatico-giuridica*, in L. Lombardi Vallauri (a cura di), *Scritti per Luigi Lombardi Vallauri*, Padova, vol. I, 2016, 503 ss.

[4] Per un'analisi dei veicoli autonomi cfr. S. Scagliarini (a cura di), Smart roads *e* driverless cars*: tra diritto, tecnologie, etica pubblica*, Torino, 2019.

[5] Al riguardo, di recente, E. Ancona (a cura di), *Soggettività, responsabilità, normatività 4.0. Profili filosofico-giuridici dell'intelligenza artificiale*, in *Rivista di Filosofia del diritto*, 1, 2019, 81-142 e A. D'Aloia (a cura di), *Intelligenza artificiale (contributi del Convegno su "Intelligenza artificiale e diritto. Come regolare un mondo nuovo", Parma, 12 ottobre 2018)*, in *BioLaw Journal*, 1, 2019, 3-182.

[6] D. Cardon, *Che cosa sognano gli algoritmi. Le nostre vite al tempo dei* big data, Milano, 2016, 5.

[7] La varietà riguarda l'eterogeneità della tipologia e dei formati dei dati, provenienti da fonti diverse; il volume si riferisce alla capacità di acquisire, memorizzare, accedere ed elaborare enormi quantità di dati; la velocità indica la capacità di acquisizione e analisi in tempo reale o ad "alta velocità".

[8] Sul tema dei *big data* cfr., *inter alia*, G. Della Morte, *Big data e protezione internazionale dei diritti umani. Regole e conflitti*, Napoli, 2018; V. Morente Parra, *Big data o el arte de analizar datos masivos. Una reflexión crítica desde los derechos fundamentales*, in *Derechos y Libertades*, 41, 2019, 225-260; A.C. Amato Mangiameli, *Algoritmi e big data. Dalla carta sulla robotica*, in *Rivista di Filosofia del diritto*, 1, 2019, 107-124; M. Delmastro, A. Nicita, *Big data. Come stanno cambiando il nostro mondo*, Bologna, 2019 e sia consentito il rinvio a F. Faini, *Data society. Governo dei dati e tutela dei diritti nell'era digitale*, Milano, 2019.

[9] Cfr. A. Mantelero, *Big data: i rischi della concentrazione del potere informativo digitale e gli strumenti di controllo*, in *Il diritto dell'informazione e dell'informatica*, 1, 2012, 135-144; G. Sartor, M. Viola De Azevedo Cunha, *Il caso Google e i rapporti regolatori USA/EU*, in *Il diritto dell'informazione e dell'informatica*, 4-5, 2014, 657-680.

alte probabilità, che permettono di fare previsioni su quello che sarà.

Di conseguenza le analisi sui *big data* permettono di estrarre preziosa conoscenza sulla realtà di riferimento e, in specifico, consentono di interpretare bisogni, profilare utenti, ottimizzare la produzione, ma anche di effettuare predizioni sui consumi, indicare preventivamente l'usura di infrastrutture e prevenire disastri, migliorare diagnosi e cure, orientare decisioni politiche[10].

In modo evidente l'utilizzo della conoscenza e delle predizioni risponde sia alla tutela di interessi generali, alla cui realizzazione sono ontologicamente orientati i soggetti pubblici, che alla realizzazione di vantaggi economici cui sono diretti i soggetti privati, in particolare multinazionali e imprese[11]. Di conseguenza, è intuitivo l'interesse che il fenomeno dei *big data* ha suscitato.

In tale contesto, gli ordinamenti giuridici sono tenuti al difficile compito di regolare tale "diluvio" di dati che caratterizza la contemporaneità e, in particolare, il diritto è chiamato a disciplinare i *big data*, al fine di tutelare i diritti dei singoli e della collettività[12].

Nonostante tale esigenza, nell'ordinamento giuridico europeo non è presente una regolazione esplicita dei *big data*, ma al riguardo rilevano due regolamenti complementari, che costituiscono il *framework* giuridico di riferimento per la circolazione libera e sicura dei dati nell'Unione europea: il regolamento europeo 2016/679 sulla protezione dei dati personali e il regolamento europeo 2018/1807 sulla libera circolazione dei dati non personali.

La distinzione tra i due regolamenti poggia sulla parallela distinzione tra i dati oggetto della disciplina europea, ossia dati personali e dati non personali.

Il regolamento (UE) 2016/679, relativo alla protezione dei dati personali, nonché alla libera circolazione di tali dati, è teso a rendere omogenea ed efficace la tutela della persona, rafforzando la correlata fiducia da parte della collettività. Seppur il regolamento europeo non tratti esplicitamente i *big data*[13], sembra consapevole dell'esistenza e delle correlate criticità: anche alla loro gestione si attagliano alcuni principi innovativi della disciplina, che mirano a un approccio proattivo e a una ponderazione preventiva dell'impatto e dei rischi sulla *data*

---

[10] Cfr., *inter alia*, D. De Pasquale, *La linea sottile tra manipolazione della rete e pubblicità*, in *Il Diritto industriale*, 6, 2012, 552 ss.; A. Mantelero, *Big data: i rischi della concentrazione del potere informativo digitale e gli strumenti di controllo*, cit., 138-139.

[11] Cfr. M. Orefice, *I big data. Regole e concorrenza*, in *Politica del diritto*, 4, 2016, 706 ss.; G. Colangelo, Big data, *piattaforme e antitrust*, in *Mercato Concorrenza Regole*, 3, 2016, 426.

[12] Sulle sfide poste alla scienza giuridica da parte delle tecnologie informatiche cfr. T. Casadei, *Presentazione. Nodi della rete e forme di regolazione*, in *Ars Interpretandi*, 1, 2017, 7-14; E. Maestri, *Lex informatica. Diritto, persona e potere nell'età del cyberspazio*, Napoli, 2015; F. De Vanna, *Diritto e nuove tecnologie: il nodo (controverso) della regolazione giuridica*, in *Lo Stato*, 11, 2018, 387-401. Per una trattazione dei rapporti tra diritto e tecnologie sia consentito il rinvio a F. Faini, S. Pietropaoli, *Scienza giuridica e tecnologie informatiche*, Torino, 2017.

[13] Cfr. l'Opinion 8/2016 «*on coherent enforcement of fundamental rights in the age of big data*», adottata il 23 settembre 2016 dall'*European Data Protection Supervisor* (EDPS) e le «*Guidelines on the protection of individuals with regard to the processing of personal data in a world of Big Data*», adottate il 23 gennaio 2017 dal Comitato consultivo della Convenzione n. 108 del 1981 sulla protezione delle persone rispetto al trattamento automatizzato di dati a carattere personale.

protection, che saranno successivamente esaminati[14].

In modo complementare, il regolamento (UE) 2018/1807 sulla libera circolazione dei dati non personali è teso a sfruttare le opportunità dell'economia basata sui dati, nella quale si collocano gli stessi *big data*. Nel rispetto dei diritti fondamentali, il regolamento prevede una serie di strumenti idonei a garantire la libera circolazione dei dati diversi dai dati personali e a migliorare la certezza giuridica e il livello di fiducia, avvalendosi anche della flessibilità necessaria a mantenerne l'efficacia nel tempo, quali il superamento degli obblighi di localizzazione dei dati, la portabilità dei dati non personali e l'autoregolamentazione affidata a codici di condotta.

Sull'interazione tra i due regolamenti rileva la comunicazione della Commissione europea «*Guidance on the Regulation on a framework for the free flow of non-personal data in the European Union*» del 29 maggio 2019, seppur sia un atto esclusivamente a titolo informativo, che per espressa dichiarazione non contiene alcuna interpretazione autorevole. Tali linee guida sono state adottate in conformità alle previsioni dell'art. 8, paragrafo 3, del regolamento europeo 2018/1807, che impone alla Commissione di pubblicare orientamenti sull'interazione tra i due regolamenti europei «in particolare per quanto concerne gli insiemi di dati composti sia da dati personali che da dati non personali». Le linee guida, al riguardo, sono consapevoli del fatto che gli insiemi di dati misti siano ricorrenti nella maggior parte delle situazioni della vita reale, soprattutto a seguito degli sviluppi dell'*Internet of Things* e dell'intelligenza artificiale.

Nel caso dei *big data* i dati sono di tipologia eterogenea e provengono da fonti diverse; di conseguenza, gli insiemi di dati misti rilevano particolarmente per l'oggetto di tale analisi, dal momento che, di norma, i *big data* avranno questa connotazione.

Nei casi in cui i dati personali e i dati non personali siano indissolubilmente legati, il regolamento (UE) 2018/1807 lascia impregiudicata l'applicazione del regolamento (UE) 2016/679[15]: secondo le linee guida, tale disposizione si traduce nell'applicazione piena del regolamento (UE) 2016/679 all'insieme di dati misti, anche quando i dati personali rappresentano soltanto una piccola parte dell'insieme dei dati. Laddove, invece, non siano indissolubilmente legati, il regolamento europeo 2018/1807 si applica alla parte dell'insieme contenente i dati non personali e, parallelamente, il regolamento europeo 2016/679 si applica alla parte dell'insieme contenente i dati personali.

Le disposizioni e le linee guida in tal modo chiariscono l'ambito di applicazione oggettivo dei rispettivi regolamenti europei nel caso di insieme di dati misti, connotazione che attiene anche al mondo dei *big data*.

[14] Cfr. G. Finocchiaro, *Introduzione al regolamento europeo sulla protezione dei dati*, in *Le Nuove leggi civili commentate*, 1, 2017, 1-18.

[15] Art. 2, paragrafo 2, reg. (UE) 2018/1807. Il concetto di "indissolubilmente legati" non è definito dai regolamenti europei, ma può essere ravvisato nelle situazioni in cui separare dati personali e dati non personali sia impossibile o ritenuto dal titolare del trattamento economicamente inefficiente o non tecnicamente realizzabile.

Di conseguenza, i *big data* sono senza dubbio toccati dai regolamenti europei sulla circolazione dei dati non personali e sulla protezione dei dati personali e, altresì, dalle linee guida, seppur rimangano sostanzialmente privi di regolazione esplicita, tranne nel caso di specifici aspetti, come quello appena esaminato afferente agli insiemi di dati misti. Ma anche sotto tale profilo l'indicazione normativa può non risultare concretamente risolutiva a causa della difficile applicazione della disciplina in materia di protezione dei dati personali ai *big data*, come sarà approfondito in seguito.

Nel caso dei *big data*, peraltro, il compito cui è chiamato il diritto è particolarmente complesso in considerazione proprio del funzionamento degli algoritmi, che evidenziano sotto diversi profili il contrasto ontologico con il modo di vedere la realtà da parte dei giuristi[16]. Seppur con una certa semplificazione, è possibile osservare che gli algoritmi prediligono un metodo descrittivo che si differenzia dal carattere prescrittivo tipico del diritto e, più ampiamente, gli algoritmi sono fondati su metodologie deterministiche, che si basano su fenomeni, circostanze oggettive, correlazioni e probabilità che rischiano di inficiare le scelte individuali e la libera volontà, su cui tipicamente poggiano gli ordinamenti giuridici democratici e la correlata regolazione[17].

In tale contrasto ontologico, il diritto si trova ad affrontare complesse questioni etico-sociali ed eterogenee problematiche giuridiche.

2. Aspetti etici e sociali

In primo luogo, è opportuno volgere lo sguardo alle questioni etiche e sociali che i *big data* sono capaci di sollevare, dal momento che tali aspetti si traducono in un possibile *vulnus* per i principi e i valori fondamentali cui il diritto si ispira.

La gestione di dati e algoritmi deve rispettare i principi dell'ordinamento, tra cui la dignità, il pieno sviluppo della persona, l'eguaglianza e la non discriminazione e, in ambito pubblico, il buon andamento dell'amministrazione pubblica fondato su imparzialità, trasparenza, pubblicità, economicità ed efficacia.

Proprio in relazione alla pubblica amministrazione sono particolarmente significative le prime interessanti sentenze amministrative relative all'utilizzo in ambito pubblico di algoritmi forieri di decisioni discutibili circa l'assegnazione della sede nei procedimenti di mobilità dei docenti nella scuola[18]. Tra queste, anche per le sentenze che ritengono ammissibile e legittimo l'utilizzo degli algoritmi in ambito pubblico, questi devono rispettare i principi generali dell'attività amministrativa, quali trasparenza, ragionevolezza e proporzionalità, devono essere conoscibili e soggetti alla cognizione, al sindacato e alla

[16] Al riguardo L. Avitabile, *Il diritto davanti all'algoritmo*, in *Rivista Italiana per le Scienze Giuridiche*, 8, 2017; A. Garapon, J. Lassègue, *Justice digitale: révolution graphique et rupture anthropologique*, Parigi, 2018.

[17] Cfr. V. Zeno-Zencovich, G. Giannone Codiglione, *Ten legal perspectives on the "Big data revolution"*, in *Concorrenza e mercato*, 2016, 49 ss.

[18] TAR Lazio, Sezione III bis, 22 marzo 2017, n. 3769; TAR Lazio, Sezione III bis, 10 settembre 2018, n. 9224; TAR Lazio, Sez. III bis, 27 maggio 2019, n. 6606; Consiglio di Stato, Sezione VI, 8 aprile 2019, n. 2270.

valutazione da parte del giudice. Di conseguenza, nel caso specifico, viene ravvisata dal Consiglio di Stato la violazione dei principi di imparzialità, pubblicità e trasparenza laddove emerga l'impossibilità di comprendere le modalità con cui, attraverso l'algoritmo, siano stati assegnati i posti disponibili[19].

In modo più esteso, involgendo il mondo pubblico e privato, nella gestione dei *big data* emerge la necessità di una valutazione dei rischi per la collettività, considerando l'impatto giuridico, sociale ed etico, al fine di prevenire i potenziali effetti negativi sulla dignità umana, sulle libertà e sui diritti fondamentali[20].

La dimensione etica è accompagnata da complessi interrogativi afferenti a quali norme etiche scegliere per educare gli algoritmi e istruire le macchine e, correlativamente, all'individuazione di principi comuni nel pluralismo etico che caratterizza la società contemporanea[21]: a tali profili si collega, altresì, la complessa valutazione relativa all'introduzione o all'adattamento di categorie giuridiche e norme.

Gli Stati gestiscono enormi quantità di dati per svolgere le proprie funzioni e i giganti tecnologici conoscono le attività, le relazioni, i pensieri e i sentimenti di tutti coloro che si servono delle loro piattaforme: rischia di prendere forma un vero e proprio *big data divide*, ossia una relazione asimmetrica tra i detentori dei *big data*, governi e colossi tecnologici, da un lato, e tutti gli altri soggetti, dall'altro[22].

I *big data* permettono di indirizzare l'individuo grazie all'analisi sulle attività digitali proprie o dei gruppi di appartenenza e di predirne i comportamenti[23]: in tale contesto possono emergere (ed emergono) profilazioni incontrollate, discriminazioni, disparità di trattamento e distorsioni del mercato. Il rischio concreto consiste nel fatto che giganti tecnologici e governi possano impiegare in vario modo le previsioni, fornite dai *big data*, anche per finalità ulteriori e diverse da quelle originarie con potenziali effetti discriminatori, minacciando principi e valori delle nostre democrazie[24].

Questa relazione asimmetrica, alimentata dal fatto che gli individui stessi, interessati alla fruizione del servizio, forniscono con estrema facilità i propri dati, può degenerare anche in forme di controllo sociale: i soggetti pubblici possono decidere di avvalersi dei dati dei colossi tecnologici e per tale via possono realizzarsi eterogenee forme di monitoraggio e sorveglianza, mettendo a rischio

[19] Consiglio di Stato, Sezione VI, 8 aprile 2019, n. 2270.

[20] In tal senso le «*Guidelines on the protection of individuals with regard to the processing of personal data in a world of Big Data*», adottate il 23 gennaio 2017 dal Comitato consultivo della Convenzione 108.

[21] A. D'Aloia, *Il diritto verso "il mondo nuovo". Le sfide dell'Intelligenza Artificiale*, in *BioLaw Journal*, 1, 2019, 3-31.

[22] M. Andrejevic, *The Big Data Divide*, in *International Journal of Communication*, 8, 2014, 1673-1689.

[23] Cfr. M.F. De Tullio, *La* privacy *e i* big data *verso una dimensione costituzionale collettiva*, in *Politica del diritto*, 4, 2016, 651 ss.

[24] Al riguardo, *inter alia*, F. Pasquale, *The Black Box Society. The Secret Algorithms That Control Money and Information*, Cambridge and London, 2015; C. O' Neil, *Armi di distruzione matematica. Come i big data aumentano la disuguaglianza e minacciano la democrazia*, trad. it., Milano, 2017.

le libertà, come nei celebri casi del *Datagate* del 2013[25] o di *Facebook – Cambridge Analytica* del 2018[26].

Il rispetto dei valori della società e dei correlati principi del diritto si traduce nella tutela che deve essere offerta ai diversi diritti in gioco, che passa dal rispetto delle norme di riferimento. Oltre a tali pericoli etico-sociali, *big data* e algoritmi, infatti, sollevano molteplici e complesse problematiche giuridiche.

3. Problematiche giuridiche

3.1. Ownership *e concorrenza*

Un primo profilo problematico di carattere giuridico afferisce alla *ownership* dei *big data*.

Risulta complesso, infatti, valutare chi ne è titolare (chi ha prodotto, chi detiene o chi elabora i dati?) e non manca chi li qualifica come beni pubblici di cui in realtà nessuno può essere considerato titolare[27].

Al riguardo si pone, altresì, la problematica afferente alla tipologia di "titolarità". Sotto tale punto di vista le raccolte di *big data* possono essere ricondotte a banche dati "non creative" e alla relativa protezione del diritto *sui generis*[28], ma permane anche il diritto d'autore dei legittimi titolari sulle informazioni strutturate e sulle banche dati creative che possono andare a formare i *big data* stessi. Anche alla luce di tali criticità di individuazione della titolarità, la tutela giuridica può essere rafforzata servendosi dell'autonomia contrattuale, al fine di tutelare i *database* in caso di cessione o nei casi di concessione di diritti di utilizzo temporaneo. Ritenendo che il valore non sia nei dati, ma nelle analisi e nelle elaborazioni algoritmiche[29], la tutela può anche concentrarsi sui relativi software, che incontrano, di norma, la protezione della proprietà intellettuale[30].

Alle problematiche di *ownership* si lega strettamente la correlata questione della responsabilità giuridica, che risulta particolarmente complessa, dal momento che si tratta di processi che spesso vedono titolarità diverse: a seconda

[25] Il cosiddetto *Datagate* ha fatto emergere la sorveglianza a livello globale sui dati personali derivante dai rapporti tra le agenzie di *intelligence* statunitensi e i colossi digitali. Al riguardo cfr. G. Greenwald, *Sotto controllo. Edward Snowden e la sorveglianza di massa*, trad. it., Milano, 2014.

[26] Lo scandalo *Facebook – Cambridge Analytica* ha riguardato l'utilizzo di dati personali per influenzare il voto in occasioni quali le elezioni negli Stati Uniti; cfr. www.ilpost.it/2018/03/19/facebook-cambridge-analytica.

[27] V. Zeno-Zencovich, *Dati, grandi dati, dati granulari e la nuova epistemologia del giurista*, in *Rivista di diritto dei media*, 2, 2018, 1 ss.

[28] Cfr. M. Falcone, *Big data e pubbliche amministrazioni: nuove prospettive per la funzione conoscitiva pubblica*, in *Rivista Trimestrale di Diritto Pubblico*, 3, 2017, 601 ss.

[29] V. Zeno-Zencovich, *Dati, grandi dati, dati granulari e la nuova epistemologia del giurista*, cit., 1 ss.

[30] Secondo V. Zeno-Zencovich, G. Giannone Codiglione, *Ten legal perspectives on the "Big data revolution"*, cit., 30 ss. le prospettive della proprietà tradizionale, della proprietà intellettuale e del contratto non sono reciprocamente esclusive, ma devono essere impiegate in base ai diversi contesti.

dei casi può cambiare la partecipazione umana all'azione che conduce a eventuali danni (si pensi all'intelligenza artificiale) e, di conseguenza, è difficile l'imputazione delle responsabilità, aprendosi anche scenari circa la personalità giuridica di nuovi soggetti[31].

In merito preme rilevare che la quantità dei dati non si traduce necessariamente in conoscenza e, di conseguenza, in valore; perché questo avvenga sono necessarie una corretta analisi, un'appropriata interpretazione e una profonda comprensione dei dati. Nel processo che dal dato conduce alla conoscenza, infatti, possono intercorrere elementi che minano il percorso, come la scarsa qualità dei dati, il loro utilizzo improprio o la manipolazione volontaria[32]: a errori e *bias* nei dati o nell'interpretazione degli stessi si affiancano anche i rischi di manipolazioni e deformazioni volontarie. In ogni caso per il diritto si pone l'esigenza di ricostruire i profili di responsabilità per rispondere ai danni che derivano da tali molteplici possibili criticità[33].

Accanto a tali profili, i *big data* sollevano questioni e problematiche di concorrenza inedite, rendendo di difficile applicazione gli istituti tipici della disciplina quali mercato rilevante, prezzi discriminatori e abuso di posizione dominante[34].

Anche sotto tale profilo, come per quello contrattuale, la differenza di rilevanza attribuita ai dati in se stessi o a quello che i dati permettono di ottenere sposta la ricostruzione giuridica relativa. Di conseguenza si contrappongono le posizioni di chi evidenzia il ruolo dei *big data*, capaci di conferire vantaggi competitivi e di erigere barriere all'ingresso, portando all'affermazione di posizioni dominanti, grazie a impedimenti legali o contrattuali alla condivisione dei dati, strategie di *lock-in* degli utenti, economie di scala ed effetti di rete diretti e indiretti, e di coloro che invece non ravvisano barriere all'entrata e che attribuiscono valore strategico ai risultati scaturenti dall'analisi dei dati e non ai dati in sé (e quindi agli algoritmi e alle tecnologie utilizzate più che ai dati stessi), che, di conseguenza, non potrebbero portare a penalizzazioni degli altri

[31] Cfr. la risoluzione del Parlamento europeo recante «*raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica*» del 16 febbraio 2017 e la comunicazione della Commissione europea «*L'intelligenza artificiale per l'Europa*» COM(2018) 237 *final* del 25 aprile 2018.

[32] Cfr. G. Colangelo, Big data, *piattaforme e antitrust*, cit., 430.

[33] I *big data* possono nutrire anche il volto oscuro delle nuove tecnologie, popolato da *fake news*, manipolazioni, discriminazioni e da ciò che compone quella che viene definita "era della post-verità". Al riguardo cfr., *inter alia*, T. Casadei (a cura di), *Post-verità*, in *Governare la paura. Journal of interdisciplinary studies*, 1, 2019, disponibile all'indirizzo governarelapaura.unibo.it/issue/view/793; C. Magnani, *Libertà di espressione e fake news, il difficile rapporto tra verità e diritto. Una prospettiva teorica*, in *Costituzionalismo.it*, 3, 2018, 1-47; G. Riva, *Interrealtà: reti fisiche e digitali e post-verità*, in *il Mulino*, 2, 2017, 326-334; M. Ferraris, *Postverità e altri enigmi*, Bologna, 2017; G. Gardini, *Le regole dell'informazione: l'era della post-verità*, Torino, 2017; G. Pitruzzella, O. Pollicino, S. Quintarelli, *Parole e potere. Libertà d'espressione, hate speech e fake news*, Milano, 2017. Per uno sguardo particolarmente pessimistico P. Savarese, *Dalla bugia alla menzogna: la postverità e l'impossibilità del diritto*, in *Nomos*, 2, 2018.

[34] Cfr. M. Orefice, *I big data. Regole e concorrenza*, cit., 722 ss. Sui rapporti tra *big data* e concorrenza, cfr. M. Maggiolino, *I big data e il diritto antitrust*, Milano, 2018.

soggetti[35].

3.2. Profili di data protection

Gli strumenti disciplinati dal regolamento (UE) 2016/679 sono tesi ad una tutela che ambisce ad essere adeguata alla società dei "grandi dati", ma *big data* e algoritmi mostrano criticità ontologiche nel rispetto della disciplina in materia di *data protection*.

I *big data* per la costante circolazione e per la continua utilizzazione cui sono sottoposti sfuggono a quel diritto di controllo su cui si fonda la protezione dei dati personali e riescono a fornire i migliori risultati proprio nelle analisi in cui non sono predefiniti gli obiettivi al momento della raccolta[36].

Tali caratteristiche rendono difficile rispettare il principio di limitazione della finalità, che prevede la raccolta dei dati personali per finalità determinate, esplicite e legittime e il successivo trattamento in modo non incompatibile con tali finalità[37]. Allo stesso modo, il volume e la varietà delle fonti dei *big data* rendono complesso il rispetto del principio di minimizzazione dei dati[38] e rischiano di inficiare la qualità, l'esattezza e l'accuratezza dei dati stessi[39].

Accanto a tali problematiche, più ampiamente nelle caratteristiche stesse di funzionamento di *big data* e algoritmi emergono criticità profonde che rischiano di minare i fondamenti stessi dell'esaminato *framework* giuridico europeo.

Il concetto di "dato personale", su cui si basa anche la dicotomia tra i regolamenti europei, può risultare insufficiente, dal momento che ci possono essere dati afferenti a gruppi o comunità, appartenenti quindi a più persone, oltre ai metadati e ai dati inferiti, estremamente significativi nel contesto dei *big data*[40]. Gli stessi dati anonimi possono non rimanere tali e le tecniche di anonimizzazione possono sollevare criticità: il pericolo sta nelle inferenze che possono essere tratte, grazie anche alla disponibilità di dati ausiliari riferibili alla persona. Ogni dato può finire per essere identificativo e quindi personale, soprattutto nelle correlazioni tra moltitudini di dati diversi[41].

Nell'universo dei "grandi dati" vacillano nella sostanza anche l'informativa da parte del titolare del trattamento[42] e il consenso libero, preventivo, specifico, inequivocabile e revocabile dell'interessato[43], dal momento che per lo più non si

[35] Cfr. G. Colangelo, Big data, *piattaforme e antitrust*, cit., 429 ss.
[36] Cfr. A. Mantelero, *Big data: i rischi della concentrazione del potere informativo digitale e gli strumenti di controllo*, cit., 135-144; G. Sartor, M. Viola De Azevedo Cunha, *Il caso Google e i rapporti regolatori USA/EU*, cit., 657-680.
[37] Art. 5, paragrafo 1, lett. b), reg. (UE) 2016/679.
[38] Art. 5, paragrafo 1, lett. c), reg. (UE) 2016/679, che prevede i criteri di adeguatezza, pertinenza e limitazione dei dati personali a quanto necessario rispetto alle finalità del trattamento.
[39] Art. 5, paragrafo 1, lett. d), reg. (UE) 2016/679.
[40] Cfr. C. Focarelli, *La privacy. Proteggere i dati personali oggi*, Bologna, 2015, 28 ss.
[41] Cfr. G. D'Acquisto, M. Naldi, *Big data e privacy by design. Anonimizzazione, pseudonimizzazione, sicurezza*, Torino, 2017, 34 ss.
[42] Artt. 12-14, reg. (UE) 2016/679.
[43] Art. 7, reg. (UE) 2016/679. Cfr. F.H. Cate, V. Mayer-Schönberger, *Notice and consent in a world of Big Data*, in *International Data Privacy Law*, vol. 3, n. 2, 2013, 67-73.

conoscono preventivamente obiettivi e finalità di utilizzo: è dubbio che in tale contesto siano rese informazioni capaci di fornire una conoscenza reale, completa e profonda e, altresì, di conseguenza, che il consenso possa considerarsi libero.

La qualificazione del consenso individuale come elemento capace di legittimare il trattamento e perfino, laddove esplicito, il processo decisionale automatizzato è particolarmente problematica[44]: il consenso preventivo, libero ed esplicito può essere ottenuto a fronte di vantaggi perseguibili e perdere così le caratteristiche che devono connotarlo[45]. Del resto la condizione dell'individuo nei rapporti con i colossi tecnologici è costituita da una libertà apparente e il mancato consenso espone all'indubbio pregiudizio di non fruire delle possibilità offerte dalle piattaforme digitali in termini relazionali e sociali: al riguardo, è particolarmente significativo il considerando 42, secondo cui «il consenso non dovrebbe essere considerato liberamente espresso se l'interessato non è in grado di operare una scelta autenticamente libera o è nell'impossibilità di rifiutare o revocare il consenso senza subire pregiudizio».

Nella consapevolezza di tali profonde criticità, che rischiano di minare le basi stesse del quadro giuridico di riferimento, alcuni principi della disciplina europea si attagliano maggiormente al mondo dei *big data*[46].

In particolare, possono essere letti sotto tale ottica i principi *data protection by design* e *by default*, previsti nell'art. 25, paragrafi 1 e 2[47], accompagnati dal *data protection impact assessment*, regolato nell'art. 35 del regolamento (UE) 2016/679[48]. Tali disposizioni prevedono che il diritto si serva della tecnica per garantire *by design* e *by default* il rispetto delle norme e danno forma a un approccio proattivo e a una valutazione preventiva dell'impatto e dei rischi dei trattamenti sulla *data protection*.

L'approccio proattivo è presente anche nell'innovativo diritto alla portabilità dei dati previsto all'art. 20 del regolamento (UE) 2016/679, idoneo a ridurre il rischio di *lock-in* e favorire la concorrenza tra piattaforme diverse[49]

[44] Artt. 7 e 22, reg. (UE) 2016/679.
[45] Cfr. G. Colangelo, Big data, *piattaforme e antitrust*, cit., 448 ss.
[46] Cfr. G. Finocchiaro, *Introduzione al regolamento europeo sulla protezione dei dati*, cit., 1-18.
[47] Il principio *data protection by design* prevede che, sia al momento di determinare i mezzi del trattamento sia all'atto del trattamento stesso, il titolare debba mettere in atto «misure tecniche e organizzative adeguate, quali la pseudonimizzazione», «volte ad attuare in modo efficace i principi di protezione dei dati, quali la minimizzazione, e a integrare nel trattamento le necessarie garanzie al fine di soddisfare i requisiti del regolamento e tutelare i diritti degli interessati», ai sensi dell'art. 25, paragrafo 1, reg. (UE) 2016/679. A tale criterio si lega il principio *data protection by default*, di cui al secondo paragrafo dell'art. 25 del reg. (UE) 2016/679: il titolare deve mettere in atto «misure tecniche e organizzative adeguate per garantire che siano trattati, per impostazione predefinita, solo i dati personali necessari per ogni specifica finalità del trattamento».
[48] Se un tipo di trattamento può presentare un rischio elevato per i diritti e le libertà delle persone fisiche, il titolare effettua, prima di procedere al trattamento, «una valutazione dell'impatto dei trattamenti previsti sulla protezione dei dati personali», ai sensi dell'art. 35, reg. (UE) 2016/679.
[49] L'interessato ha il diritto di ricevere in un formato strutturato, di uso comune e leggibile da dispositivo automatico, i dati personali che lo riguardano forniti a un titolare e di

(principio presente, seppur in modo diverso, anche nel regolamento europeo 2018/1807), e in altri strumenti, come la consultazione preventiva (art. 36)[50] e il *data breach* (artt. 33-34)[51].

Nel regolamento europeo 2016/679, in linea con tale approccio di tutela si pongono la logica di *accountability* e responsabilizzazione dei soggetti che trattano i dati personali (art. 24), la figura del *Data Protection Officer* (DPO) (artt. 37-39)[52] e la previsione della contitolarità, accompagnata dalla definizione delle rispettive responsabilità (art. 26), disposizioni coadiuvate sia dall'attenzione alla sicurezza (art. 32), sia dall'effettività e dall'efficacia del sistema sanzionatorio (artt. 82-84). Tali norme si attagliano in modo efficace alle caratteristiche dei *big data* e alla filiera di soggetti diversi che spesso ne caratterizza la gestione.

Ai fini di tale analisi rileva particolarmente l'art. 22 del regolamento (UE) 2016/679, dedicato al «processo decisionale automatizzato relativo alle persone fisiche, compresa la profilazione», che pare alludere chiaramente ai *big data*: «l'interessato ha 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».

La disposizione, però, non si applica laddove la decisione sia necessaria per la conclusione o l'esecuzione di un contratto, sia autorizzata dal diritto europeo o nazionale o si basi sul consenso esplicito dell'interessato, che anche in tale contesto appare come strumento capace di legittimare il trattamento (strumento in realtà improprio, come sopra esaminato). Tali eccezioni sono bilanciate obbligando il titolare del trattamento, nei casi di consenso e conclusione o esecuzione del contratto, ad attuare comunque «misure appropriate per tutelare i diritti, le libertà e i legittimi interessi dell'interessato, almeno il diritto di ottenere l'intervento umano da parte del titolare del trattamento, di esprimere la propria opinione e di contestare la decisione».

4. Principi e strumenti della *data governance*: tecnologia, trasparenza e controllo collettivo

Al fine di costruire una solida *data governance* ispirata ai principi etici e ai valori degli ordinamenti democratici, è necessario superare le problematiche esaminate, che a guardare bene trovano fondamento nella chiusura dei processi di gestione dei dati, nel significativo squilibrio tra le parti e nella conseguente inevitabile

trasmettere tali dati a un altro titolare, senza impedimenti da parte del primo titolare cui li ha forniti.

[50] Il titolare, prima di procedere al trattamento, consulta l'autorità di controllo qualora la valutazione d'impatto indichi che il trattamento presenterebbe un rischio elevato in assenza di misure adottate per attenuare il rischio.

[51] Il titolare ha l'obbligo di notificare eventuali violazioni dei dati personali all'autorità nazionale nei tempi e nelle modalità previste e, se la violazione è suscettibile di presentare un rischio elevato per i diritti e le libertà delle persone fisiche, deve comunicare la violazione all'interessato senza ingiustificato ritardo.

[52] Si tratta di una figura prevista con la funzione di garantire una corretta gestione dei dati.

incapacità del singolo di potersi tutelare in modo efficace.

In primo luogo, uno strumento per tutelare i diritti, riducendo chiusura e opacità nella gestione dei dati, può risiedere proprio nella stessa tecnologia. Come esaminato, nella disciplina in materia di *data protection* il diritto si avvale della tecnica per assicurare il suo rispetto e tutelare la persona per mezzo dei principi *data protection by default* e *by design* e di strumenti come il *data protection impact assessment*: la tecnologia si pone quale rimedio preventivo a possibili violazioni delle norme, proteggendo la persona fin dalla progettazione, per impostazione preventiva e per mezzo della valutazione d'impatto. Tale approccio può essere impiegato non solo per il rispetto della disciplina in materia di protezione dei dati personali, ma può essere utilizzato più ampiamente per affrontare le problematiche giuridiche coinvolte nella gestione degli insiemi di dati misti e, quindi, dei *big data*: il diritto può avvalersi della tecnica per far rispettare i suoi principi di riferimento.

In secondo luogo, per contrastare l'opacità nell'utilizzo dei *big data*, insieme alla tecnologia, è necessario affidarsi a una trasparenza sostanziale, che per essere tale deve tradursi in un dovere di lealtà dei titolari del trattamento nei confronti degli interessati. Anche in tal caso risulta significativo quanto previsto dalla disciplina in materia di *data protection*: il titolare del trattamento è tenuto a fornire all'interessato, tra le informazioni necessarie per garantire un trattamento corretto e trasparente, «l'esistenza di un processo decisionale automatizzato, compresa la profilazione [...] e, almeno in tali casi, informazioni significative sulla logica utilizzata, nonché l'importanza e le conseguenze previste di tale trattamento per l'interessato»[53]. Su tali informazioni esiste, altresì, un diritto di accesso riconosciuto all'interessato esplicitamente dall'art. 15 del regolamento (UE) 2016/679.

Nel caso dei *big data*, che si avvalgono di processi decisionali automatizzati, la norma si traduce nella necessità di fornire informazioni e garantire l'accesso in merito alla logica utilizzata dagli algoritmi, all'impatto e alle conseguenze per l'interessato. Tale forma di trasparenza permette di confinare o quanto meno di rilevare errori, *bias*, manipolazioni che, come esaminato, possono inficiare i *big data*. Inoltre in tal modo è possibile riequilibrare l'asimmetria tra le parti, imponendo maggiore trasparenza in merito agli algoritmi stessi, al fine di garantire la consapevole autodeterminazione e la correlata libertà degli individui.

Con questo approccio si coniugano l'apertura e il rilascio dei dati in *open data*, esigenze che emergono in modo evidente nella normativa europea relativa ai dati pubblici[54]. Insieme alla trasparenza, l'apertura dei *big data* in ambito pubblico, ma anche in ambito privato potrebbe contribuire a sanare le asimmetrie informative, mettendo a disposizione della collettività i dati, anche se inevitabilmente provocherebbe una perdita di potere per i detentori degli stessi.

[53] Art. 13, paragrafo 2, lett. f) e art. 14, paragrafo 2, lett. g), regolamento (UE) 2016/679.

[54] Gli *open data* sono i dati resi disponibili con le caratteristiche tecniche e legali necessarie per essere liberamente utilizzati, riutilizzati e ridistribuiti da chiunque, in qualsiasi momento e ovunque; cfr. *opendatacharter.net*.

Tale ricostruzione, peraltro, si attaglia alla qualificazione dei dati come "beni pubblici", dei quali possiedono quelle caratteristiche di non esclusività e non rivalità, al netto di concrete pretese di appropriazione e restrizioni surrettizie capaci di renderli di fatto "proprietà" di pochi soggetti.

Infine il significativo squilibrio tra le parti in gioco si traduce nel rischio di "solitudine" del singolo, spesso inconsapevole e in ogni caso incapace di potersi tutelare in modo efficace da solo. Al riguardo, i valori della dignità, dello sviluppo della persona e dell'uguaglianza a fronte delle possibili manipolazioni e discriminazioni, che gli algoritmi sono capaci di realizzare, interessano tutta la collettività; parimenti, di conseguenza, può essere opportuno immaginare una tutela collettiva dell'individuo, che si affianchi a quella individuale[55].

Un riferimento significativo è nella disciplina in materia di protezione dei dati personali e, in specifico, nell'art. 80 del regolamento europeo 2016/679, secondo paragrafo: «Gli Stati membri possono prevedere che un organismo, organizzazione o associazione [...], indipendentemente dal mandato conferito dall'interessato, abbia il diritto di proporre, in tale Stato membro, un reclamo all'autorità di controllo competente, e di esercitare i diritti di cui agli articoli 78 e 79[56], qualora ritenga che i diritti di cui un interessato gode a norma del [...] regolamento siano stati violati in seguito al trattamento». In tal modo la protezione della persona si può liberare dalla necessità di azione da parte dell'individuo stesso e può essere attivata da organizzazioni e associazioni nel momento in cui siano violati quei "beni pubblici" e quei valori collettivi protetti dalla disciplina.

Gli strumenti esaminati, afferenti a tecnologia, trasparenza, apertura e tutela collettiva, trovano collante comune nell'esigenza di dare particolare rilievo nella *data governance* alla dimensione etica.

I *big data* necessitano di un uso etico, consapevole e socialmente responsabile dei dati, idoneo a valutare la possibilità di conflitto con altri diritti e il rispetto dei principi di limitazione delle finalità e di trasparenza, in modo da evitare che i dati siano ulteriormente elaborati in modo inaspettato, inappropriato o discutibile per l'interessato. Al riguardo emerge un diritto di controllo sui dati non circoscritto all'individuo, ma tale da comprendere una valutazione dei rischi per la collettività, considerando l'impatto giuridico, sociale ed etico dell'utilizzo dei *big data* sia a livello individuale che collettivo, al fine di prevenire i potenziali effetti negativi sulla dignità umana, sulle libertà e sui diritti fondamentali[57].

---

[55] Cfr. A. Mantelero, *I Big Data nel quadro della disciplina europea della tutela dei dati personali*, in *Il Corriere giuridico*, ed. speciale, 2018, 54 ss.

[56] Si tratta del diritto a un ricorso giurisdizionale effettivo nei confronti dell'autorità di controllo e del diritto a un ricorso giurisdizionale effettivo nei confronti del titolare o del responsabile del trattamento.

[57] In tal senso le «*Guidelines on the protection of individuals with regard to the processing of personal data in a world of Big Data*», adottate il 23 gennaio 2017 dal Comitato consultivo della Convenzione 108.

ISSN: 2037-6677

La gestione etica, consapevole e sociale dei dati, basata sulla trasparenza e orientata a una tutela collettiva accanto a quella individuale, rappresenta la strada su cui è possibile immaginare la *data governance* nel nostro futuro.