L'IMPATTO DEI DATI NON PERSONALI SULLE DECISIONI ALGORITMICHE: LA PROSPETTIVA DELLE AUTORITÀ AMMINISTRATIVE INDIPENDENTI EUROPEE*

STEFANO TORREGIANI**

Sommario

Abstract

The ever-expanding power of algorithms designed to make decisions that significantly affect physical persons raises important questions regarding the way the law should protect individuals. The hegemonic capacity proven by the regulation on the processing of personal data, always considered the most suitable basis to legally frame automated decisions, has partly eclipsed other issues and legislations. This paper aims to examine the phenomenon of algorithmic decision-making through the lens of non-personal data by analyzing the documents of some of the European independent administrative authorities that focus on the issues stemming from the lacking attention paid to the strict relation between this type of data and algorithmic systems. Lastly, this work tries to detect the gaps of the legal framework and to understand how the European legislator could address the topic of the algorithmic decision-making in the next years.

Suggerimento di citazione

S. TORREGIANI, *L'impatto dei dati non personali sulle decisioni algoritmiche: la prospettiva delle autorità amministrative indipendenti europee,* in *Osservatorio sulle fonti,* n. 2/2021. Disponibile in: http://www.osservatoriosullefonti.it

* Il contributo costituisce la rielaborazione della relazione tenuta al *webinar* "Autorità amministrative indipendenti e regolazione delle decisioni algoritmiche" svoltosi il 7 maggio 2021 e organizzato dal Dipartimento di Scienze Giuridiche dell'Università di Firenze, nell'ambito del Progetto PRIN 2017 *Self- and Co-regulation for Emerging Technologies: Towards a Technological Rule of Law* (SE.CO.R.E TECH).

** Dottorando di ricerca in Scienze giuridiche presso l'Università degli Studi di Macerata.
Contatto: s.torregiani@unimc.it

Registrazione presso il Tribunale di Firenze n. 5626 del 24 dicembre 2007
Rivista inclusa nella classe A delle Riviste scientifiche di Area 12 – Direttore Prof. Paolo Caretti
ISSN 2038-5633

1. Introduzione

La possibilità di demandare ad un'entità impersonale la facoltà di prendere decisioni capaci di incidere sulla sfera giuridica degli individui costituisce uno dei tratti caratteristici dell'era cibernetica,[1] dove la tecnologia e l'estrema fiducia che riponiamo in essa stanno relegando l'essere umano ad un ruolo sempre più marginale. In questo contesto, il tema della decisione algoritmica ha attirato l'interesse di giuristi spinti dal desiderio di trovare il corretto inquadramento giuridico dell'algoritmo. La poliedricità del fenomeno ha sollecitato a più riprese anche l'intervento di autorità amministrative indipendenti di ogni ordine e grado, costrette ad affrontare una sfida per la quale gli strumenti a disposizione non sempre si sono rivelati efficaci.

Il fatto che la disciplina relativa al trattamento dei dati personali sia sempre stata considerata – a ragione –[2] come la base più solida per fornire una risposta alle questioni sollevate dall'attuazione delle decisioni automatizzate ha in parte eclissato la rilevanza che fattispecie o discipline diverse potrebbero assumere in merito. Pertanto, il presente contributo mira ad esaminare il fenomeno della decisione algoritmica attraverso il prisma dei dati non personali, fattispecie opposta e residuale rispetto a quella dei dati personali, analizzando alcuni fra i documenti delle autorità amministrative indipendenti, nazionali e continentali, che hanno messo in risalto le questioni di diritto poste dalla normativa europea sui dati con riguardo al complesso mondo degli algoritmi.

Non essendo stato istituito un organo di controllo indipendente *ad hoc* dalla disciplina sui dati non personali recentemente varata,[3] i documenti selezionati consistono in opinioni o studi condotti da autorità indipendenti con competenze in settori come la protezione dei dati, la concorrenza o la comunicazione in cui le informazioni a carattere non personale svolgono una funzione chiave, pur rimanendo sullo sfondo. Al di là della forma, il contenuto dei testi proposti si rivela estremamente prezioso perché capace di mettere in luce in maniera ancor più chiara le lacune del nostro ordinamento in tema di decisioni algoritmiche e di offrire qualche suggerimento sulla strada che il legislatore potrebbe intraprendere per il prossimo futuro.

[1] A. SIMONCINI, *L'algoritmo incostituzionale: intelligenza artificiale e il futuro delle libertà*, in *BioLaw Journal – Rivista di BioDiritto*, 2019, 63 ss.

[2] Per un approfondimento sulle disposizioni del Regolamento sulla protezione dei dati personali che regolano i processi decisionali automatizzati e la profilazione si veda: Article 29 Data Protection Working Party (WP29), "G*uidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679*" (WP251).

[3] Il Regolamento (UE) 2018/1807, relativo a un quadro applicabile alla libera circolazione dei dati non personali nell'Unione europea, attribuisce compiti di monitoraggio del rispetto della disciplina principalmente alla Commissione europea.

2. Premessa: le principali questioni riguardanti il dato non personale

Prima di addentrarsi nell'analisi dei testi rilevanti, è opportuno delineare, per sommi capi, la fattispecie del dato non personale in linea con quanto stabilito dal quadro europeo e le principali questioni giuridiche ad esso collegate.

Dal punto di vista normativo, la prima fonte generale che ha fornito una definizione positiva è stata il Regolamento (UE) 2018/1807, relativo a un quadro applicabile alla libera circolazione dei dati non personali nell'Unione europea.[4] Sebbene speculari e residuali rispetto al GDPR,[5] le disposizioni che lo compongono non hanno predisposto un sistema paragonabile a quello in vigore per i dati personali, giacché mirano principalmente a favorire la libera circolazione delle informazioni non attinenti a persone fisiche per mezzo dell'imposizione di un divieto per gli Stati membri di introdurre o mantenere obblighi di localizzazione, ossia misure normative o prassi limitanti per il flusso di dati all'interno del territorio europeo.[6] Al di là di tale previsione – che costituisce il pilastro dell'intero impianto – il Regolamento è divenuto celebre principalmente per la definizione di dato non personale introdotta nell'ordinamento: in tale categoria rientrano tutti "i dati diversi dai dati personali definiti all'articolo 4, punto 1, del regolamento (UE) 2016/679".[7] La nozione fissata dal legislatore colpisce per la sua indeterminatezza e per le problematiche che ne conseguono, in quanto presuppone che l'altra categoria, quella dei dati personali, sia esente da difficoltà di tipo definitorio. Al contrario, come dimostrato anche dalla Corte di giustizia dell'Unione europea,[8] l'accertamento della qualifica personale dei dati è un compito tutt'altro che semplice, poiché ognuno di essi è composto da più elementi che concorrono congiuntamente a determinarne la natura.[9] Sicché, la scelta di optare per una definizione a carattere

[4] Regolamento (UE) 2018/1807 del Parlamento europeo e del Consiglio, del 14 novembre 2018, relativo a un quadro applicabile alla libera circolazione dei dati non personali nell'Unione europea.

[5] 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.

[6] Art. 4, par. 1, Regolamento (UE) 2018/1807: "Gli obblighi di localizzazione di dati sono vietati a meno che siano giustificati da motivi di sicurezza pubblica nel rispetto del principio di proporzionalità".

[7] Art. 3, punto 1, Regolamento (UE) 2018/1807.

[8] Corte di Giustizia UE, 19-10-2016, causa C-582/14, Patrick Breyer contro Bundesrepublik Deutschland.

[9] Il Gruppo di lavoro articolo 29 già nella vigenza della precedente disciplina aveva affermato che per stabilire se un dato debba considerarsi come personale è necessario prendere in considerazione ognuno dei quattro elementi indicati nella definizione: "qualsiasi informazione", "concernente", "identificata o identificabile", "persona fisica". Al riguardo, si veda il "Parere 4/2007 sul concetto di dati personali" (WP136).

negativo non può che tradursi in un difetto a monte del sistema, destinato poi a riverberarsi sull'intera disciplina dei dati, sia personali che non.[10]

La superficialità definitoria del testo del regolamento viene tamponata solo in parte dalla *Guidance* pubblicata dalla Commissione europea che, tramite alcuni chiarimenti interpretativi, aveva lo scopo di favorirne l'applicazione.[11] Difatti, la guida cerca di dare maggiore corpo alla nozione rispetto a quanto fatto dalle disposizioni normative,[12] e suddivide la fattispecie del dato non personale in due sub-categorie: da un lato, i dati che non sono mai stati riferiti ad una persona identificabile, spesso definiti "industriali" poiché attinenti a parametri tipici dell'industria o dell'ambiente; e, dall'altro, i dati anonimizzati, ossia quelli dai quali sono stati rimossi, attraverso una procedura di anonimizzazione, i fattori che permettevano l'identificazione del soggetto cui originariamente si riferivano. Seppur apprezzabile nello scopo, il lavoro della Commissione si arresta alla semplice menzione di queste fattispecie, senza scendere nel dettaglio delle questioni giuridiche che le concernono.

Per i dati anonimizzati, le problematiche nascono nella fase di transizione che porta al mutamento della loro natura, da personale a non personale, poiché possono residuare alcune "tracce di personalità" che aumentano il rischio che la persona fisica possa essere re-identificata. Al considerando 26, il GDPR esclude questo tipo di dati dal suo ambito di applicazione, ma, se si eccettua un generico riferimento ai mezzi di cui il titolare del trattamento o un terzo possono ragionevolmente avvalersi per identificare direttamente o indirettamente l'individuo, nulla prevede con riguardo alle procedure che concretamente possono garantire una corretta anonimizzazione.[13] La visione europea sul tema è ancora cristallizzata nell'opinione del Gruppo di lavoro articolo 29 che già a pochi anni di distanza dalla sua pubblicazione si è rivelata, in parte, anacronistica.[14] Il principio della irreversibilità assoluta della anonimizzazione

[10] Parere del Comitato economico e sociale europeo (CESE) sulla «Proposta di regolamento del Parlamento europeo e del Consiglio relativo a un quadro applicabile alla libera circolazione dei dati non personali nell'Unione europea» (2018/C 227/12).

[11] Comunicazione della Commissione al Parlamento europeo e al Consiglio, *Guidance on the Regulation on a framework for the free flow of non-personal data in the European Union* (COM (2019) 250 final).

[12] Il considerando n. 9 del Regolamento (UE) 2018/1807 accenna solo brevemente ad alcune ipotesi di dati non personali senza procedere alla loro classificazione.

[13] Considerando n. 26, Regolamento (UE) 2016/679.

[14] Article 29 Data Protection Working Party (WP29), "Opinion 05/2014 on Anonymisation Techniques" (WP216). In questa opinione, il Gruppo afferma che una corretta procedura di anonimizzazione può considerarsi tale solamente se garantisce l'irreversibilità dell'operazione. È opportuno sottolineare che l'opinione del Gruppo di lavoro articolo 29 è stata espressa nel contesto della precedente direttiva ed in riferimento alle definizioni in essa contenute; a tale riguardo, si veda: R. HU, S. STALLA-BOURDILLON, M. YANG, V. SCHIAVO, V. SASSONE, *Bridging Policy, Regulation and Practice? A techno-legal Analysis of Three Types of Data in the GDPR*, in R. LEENES, R. VAN BRAKEL, S.

che traspare dal documento del Gruppo non sembra più appartenere all'epoca dei Big Data. I moderni algoritmi e la crescente disponibilità di dataset hanno aumentato considerevolmente le possibilità di re-identificare – anche involontariamente – le persone fisiche, rendendo quasi impossibile garantire con assoluta certezza l'irreversibilità del processo. Per tali ragioni, buona parte della dottrina spinge per un approccio più dinamico all'anonimizzazione, in ragione del quale si accetta che il rischio di re-identificazione non possa in nessun caso essere ridotto a zero e si stabilisce una soglia giuridica al di sotto della quale esso possa considerarsi tollerabile.[15]

L'altro punto che difetta di adeguato approfondimento da parte del legislatore, ma che già da tempo anima il dibattito tra giuristi ed economisti, concerne l'introduzione di un nuovo diritto di proprietà sui dati. La raccolta perpetua e massiva effettuata da onnipresenti sensori intelligenti ha aggiunto ulteriore complessità nella ricerca di uno strumento normativo capace di fornire garanzie adeguate a coloro che trattano le informazioni nel mondo digitale: se le caratteristiche tipiche dei dati, segnatamente, l'immaterialità, la non rivalità e l'escludibilità *de facto*, avevano già messo in crisi le tradizionali categorie giuridiche relative ai beni, ora appare ancora più arduo ricorrere a discipline di settore come quelle a tutela delle banche dati o dei segreti commerciali, il cui ambito di applicazione non sembra combaciare con il moderno sistema di raccolta ed elaborazione dei dati.[16] Tuttavia, al momento, tanto a livello istituzionale quanto dottrinale, sta acquisendo sempre maggiore credito la teoria secondo cui la previsione di un vero e proprio diritto di proprietà, inizialmente proposto allo scopo di offrire maggiori garanzie ai soggetti deboli coinvolti nella catena di trattamento dei dati,[17] si tramuterebbe in un indebito ostacolo alla libera circolazione delle informazioni e all'avanzamento della tecnologia.[18]

Un ultimo profilo che merita particolare attenzione per la sua intima connessione con il tema della decisione algoritmica riguarda gli insiemi di dati misti, ossia composti da dati sia personali che non personali. A tale riguardo, tanto il Regolamento 2018/1807 quanto la guida della Commissione riflettono in

GUTWIRTH, P. DE HERT (a cura di), *Data Protection and Privacy: The Age of Intelligent Machines*, Bloomsbury Publishing, 2017, 119 ss.

[15] S. STALLA-BOURDILLON-KNIGHT, *Anonymous data v. Personal Data – A false debate: an EU perspective on anonymization, pseudonymization and personal data*, in *Wisconsin International Law Journal*, 2017, 284 ss.

[16] J. DREXL, *Designing Competitive Markets for Industrial Data – Between Propertisation and Access*, in *Max Planck Institute for Innovation & Competition Research Paper No. 16-13*, 2016, 1 ss.

[17] A. WIEBE, *Protection of industrial data – a new property right for the digital economy?*, in *Journal of Intellectual Property Law & Practice*, 2016, 1 ss.

[18] W. KERBER, *A New (Intellectual) Property Right for Non-Personal Data? An Economic Analysis*, *Joint Discussion Paper Series in Economics by the Universities of Aachen, Gießen, Göttingen, Kassel, Marburg, Siegen*, 2016.

maniera evidente il tipico approccio europeo che pervade l'intero ordinamento giuridico in materia di dati. L'articolo 2, paragrafo 2, dopo aver disposto che ogni categoria di dato ricade nella disciplina per esso prevista – inserendo, di fatto, un onere nascosto in capo al titolare di preventiva identificazione e successiva scissione dell'insieme –,[19] stabilisce che nelle ipotesi in cui gli insiemi siano "indissolubilmente legati", ossia che separare il dataset in due parti, una composta di dati personali e l'altra di dati non personali, risulterebbe impossibile o economicamente svantaggioso, si applicano le disposizioni del GDPR all'intero insieme, anche nel caso in cui i dati a carattere personale siano presenti solo in minima parte.[20] Una scelta così sbilanciata non fa altro che sminuire il valore del dato non personale, incentivando un approccio miope che getta ombra su una fattispecie fin troppo negletta nel panorama continentale.

In sostanza, si manifesta in modo evidente la tradizionale impostazione europea che pone il dato personale al centro della regolamentazione del mondo dei dati e che, per converso, non permette né di apprezzare pienamente la categoria del dato non personale, né di concentrarsi debitamente nella ricerca di soluzioni adeguate alle questioni anzidette. Ne deriva un ordinamento giuridico a doppia velocità dove due fattispecie che si pongono in continuità l'una con l'altra e che si influenzano vicendevolmente permangono in una situazione di squilibrio regolatorio che, in un momento storico dove l'utilizzo degli algoritmi ha gettato maggiore ombra sulla linea di confine che separa i dati non personali da quelli personali,[21] rischia di tradursi in un meccanismo inadeguato a far fronte alla complessità della società digitale.

3. Gli interventi delle autorità amministrative indipendenti

Alla luce di tale premessa, è possibile apprezzare la rilevanza dell'impatto che la regolamentazione bifronte predisposta dal legislatore continentale è in grado di avere in un contesto come quello dei sistemi algoritmici, dove i set di dati alla base della decisione finale racchiudono informazioni di svariata natura e provenienza. I documenti delle autorità amministrative indipendenti analizzati nel corso del presente lavoro sono stati prescelti proprio per la loro capacità di mettere in risalto le difficoltà che sorgono dinnanzi a questa aporia ordinamentale e per le soluzioni, *de jure condito* e *de jure condendo*, che sono state proposte per superarla. La sequenza con cui verranno presentati restituisce uno

[19] A tale proposito, sia consentito rinviare a: S. TORREGIANI, *Il dato non personale alla luce del Regolamento (UE). 2018/1807: tra anonimizzazione, ownership e Data by Design*, in *federalismi.it*, 2020, 317 ss.

[20] Comunicazione della Commissione, *Guidance on the Regulation on a framework for the free flow of non-personal data in the European Union,* cit.

[21] S. FORGE, *Optimal Scope for Free-Flow of Non-Personal Data in Europe*, Briefing Requested by the IMCO committee, 2016.

spaccato interessante dell'evoluzione del contesto in cui operano i soggetti deputati al controllo del corretto utilizzo delle nuove tecnologie e riflette una crescente insofferenza di fronte alla parziale incompletezza dell'ordinamento giuridico attualmente vigente nell'Unione europea in riferimento alle dinamiche algoritmiche.

3.1 Il primo argine: l'interpretazione del concetto di identificabilità

Il primo documento che si intende presentare riproduce in maniera emblematica il modo in cui le autorità amministrative, ed in generale le istituzioni europee, affrontano i problemi di qualificazione dei dati nelle ipotesi in cui questa non dovesse risultare di immediata individuazione. A tale proposito, la disciplina continentale contempla uno strumento che allarga notevolmente la capacità di interpretazione estensiva degli organi di controllo: il concetto di "identificabilità".[22] Malgrado la maggiore complessità definitoria inaugurata dal GDPR, il quale ha aggiunto ulteriori sfumature alla nozione di dato personale, l'approccio binario presente nella vigenza della precedente disciplina costituisce ancora la struttura portante dell'impianto europeo: le informazioni che possono essere collegate, anche in via indiretta, ad una persona fisica sottostanno alle disposizioni del GDPR, indipendentemente dalla difficoltà e dal numero di passaggi necessari a risalire all'identità del *data subject*. In virtù di tale assunto, i dati che si riferiscono alla persona identificata e quelli che riguardano la persona identificabile sono equiparati dal punto di vista giuridico, nonostante nel secondo caso sia comunque presente un certo grado di de-identificazione.[23] Di fronte alla alternativa secca predisposta dal legislatore, l'interprete si è sempre schierato in favore di una lettura estensiva che ha nel tempo allargato notevolmente l'area del dato personale, estendendo le garanzie accordate all'individuo anche a scenari in cui il rischio di lesione del diritto fondamentale alla protezione delle proprie informazioni personali risulti minimo, se non addirittura nullo.[24]

Orbene, se l'identificabilità ricorre in tutte quelle ipotesi in cui il legame, non ancora noto, tra la persona fisica e l'informazione può essere scoperto

[22] Art. 4, punto 1), GDPR: "«dato personale»: qualsiasi informazione riguardante una persona fisica identificata o identificabile («interessato»); si considera identificabile la persona fisica che può essere identificata, direttamente o indirettamente, con particolare riferimento a un identificativo come il nome, un numero di identificazione, dati relativi all'ubicazione, un identificativo online o a uno o più elementi caratteristici della sua identità fisica, fisiologica, genetica, psichica, economica, culturale o sociale".

[23] M. HINTZE, *Viewing the GDPR through a de-identification lens: a tool for compliance, clarification, and consistency*, in *International Data Privacy Law*, Volume 8, Issue 1, 2018, 86 ss.

[24] C. FOGLIA, *Il dilemma (ancora aperto) dell'anonimizzazione e il ruolo della pseudonimizzazione nel GDPR*, in PANETTA R. (a cura di), *Circolazione e protezione dei dati personali tra libertà e regole del mercato. Commentario al Regolamento UE n. 679/2016 e al d.lgs. n. 101/2018*, 2019, 309 ss.

tramite i mezzi di cui il titolare del trattamento, o un terzo, può ragionevolmente avvalersi,[25] è evidente che nelle situazioni limite si amplificano le possibilità di dilatare in via esegetica il campo applicativo del regolamento. Quanto detto vale, *a fortiori*, nell'epoca della raccolta massiva ed indiscriminata di dati effettuata tramite sensori incorporati in dispositivi intelligenti, in virtù della quale ad informazioni di evidente carattere personale se ne aggiungono di ulteriori che non concernono in via diretta le persone fisiche, ma che, potenzialmente, potrebbero permettere l'identificazione con l'aiuto di altri dati in possesso del medesimo titolare del trattamento.[26]

In tal senso, il settore delle auto connesse, caratterizzato da intense attività di raccolta, elaborazione e trasmissione di dati di varia natura, offre una panoramica interessante del rilievo che assume lo strumento dell'identificabilità nelle mani del soggetto pubblico nell'assolvimento dei compiti di controllo del soggetto privato. Sin dalle prime fasi di sviluppo, le associazioni di settore che rappresentano i più importanti produttori di automobili in Europa si sono cimentate in una non semplice operazione di qualificazione dei dati trattati dai sensori integrati nei veicoli connessi.[27] Al netto di piccole differenze tassonomiche, tutti i documenti dimostrano particolare attenzione alle istanze relative alla tutela del diritto alla protezione dei dati personali e, per tali ragioni, sottolineano come questa categoria di dati possa comprendere tutte quelle informazioni che tramite altri indicatori permettono di risalire all'identità degli occupanti il veicolo. Tuttavia, dopo tale premessa, la classificazione degli altri dati, quelli non direttamente collegati con alcuna persona fisica ma la cui elaborazione è comunque necessaria a fornire i servizi generalmente previsti nel contratto di compravendita dell'automobile, sembra rispondere a logiche di ordine squisitamente tecnico, dove all'attenzione per la protezione dei dati personali fanno da contraltare interessi di carattere eminentemente industriale, come la proprietà, la segretezza e la sicurezza.[28]

Una diversa sensibilità giuridica si percepisce, invece, nella risposta proveniente dal lato pubblico, giacché gli organi di controllo preferiscono

[25] P. SCHWARTZ, D. SOLOVE, *The PII Problem: Privacy and a New Concept of Personally Identifiable Information*, in *New York University Law Review*, Vol. 86, 2011, 1814 ss.

[26] I. RUBINSTEIN, W. HARTZOG, *Anonymization and Risk*, in *New York University Public Law and Legal Theory Working Papers*, 2015, 703 ss.

[27] Verband der Automobilindustrie (VDA), *Access to the vehicle (and vehicle generated data)*, 2016, disponibile all'indirizzo: https://www.vda.de/en/topics/innovation-and-technology/network/access-to-the-vehicle.html; Society of Motor Manufacturers and Traders (SMMT), *Connected and Autonomous Vehicles – Position Paper*, February 2017, disponibile all'indirizzo: https://www.smmt.co.uk/wp-content/uploads/sites/2/SMMT-CAV-position-paper-final.pdf.

[28] Osborne Clarke Rechtsanwälte Steuerberater Partnerschaft mbB, *What EU legislation says about car data. Legal Memorandum on connected vehicles and data*, for Fédération Internationale de l'Automobile, 2017.

abbracciare una lettura che si allontana dalla presunzione relativa di non personalità del dato tecnico prospettata dalle associazioni di settore per determinati tipi di informazione. In tal senso, l'autorità di protezione dei dati personali della Repubblica federale tedesca (Bundesbeauftragter für den Datenschutz und die Informationsfreiheit, BfDI)[29] e quella della Repubblica francese (Commission nationale de l'informatique et des libertés, CNIL)[30] prescindono dalla medesima "mappatura tecnica", per concentrarsi quasi esclusivamente sui profili di identificabilità che caratterizzano i dati trattati dal veicolo: la facilità con cui i titolari del trattamento possono collegare i dati tecnici al numero identificativo di chi possiede o conduce il veicolo non permette di addivenire ad una classificazione *ex ante* così netta, poiché si rischia di trarre in inganno e di condurre a semplificazioni aprioristiche che portano alla fuoriuscita dal perimetro del GDPR di informazioni che potrebbero permettere l'identificazione del *data subject* in maniera rapida e poco difficoltosa. A fronte di tali considerazioni, le autorità tendono probabilmente a sovrastimare il rischio che l'identità delle persone fisiche, sia a bordo che all'esterno del veicolo, possa essere svelata e, pertanto, estendono le tutele predisposte dalla disciplina dei dati a carattere personale anche a quei dati, ad esempio relativi all'usura o ai malfunzionamenti, che a prima vista potrebbero sembrare esclusivamente tecnici.

Questo meccanismo di difesa di stampo europeo tende ad espandere l'area del giuridicamente rilevante per diminuire i rischi di lesione della sfera individuale, ma, allo stesso tempo, comporta un aumento degli oneri gravanti sul titolare del trattamento e limita l'utilizzo di dati come *input* degli algoritmi, causando di fatto un rallentamento nell'analisi di dati e nella crescita industriale. Se per un verso questa scelta potrebbe far sentire più al sicuro il cittadino europeo, per un altro, le imprese continentali ne escono enormemente danneggiate rispetto alle concorrenti extracomunitarie. In tal senso, un termine di paragone interessante viene offerto dall'ordinamento statunitense, dove non è presente una tutela costituzionale dell'*informational privacy* come in Europa e dove la logica che informa la protezione dei dati si ispira a ragioni di tipo consumeristico.[31] Anche oltreoceano si è verificato un dibattito simile, benché in un settore parzialmente diverso, nell'ambito del quale ai cosiddetti *agricultural technology provider*, ossia le imprese che utilizzano gli algoritmi per analizzare dati agricoli e fornire indicazioni relative alla coltivazione, si

[29] *Joint statement of the conference of the independent data protection authorities of the Federal and State Governments of Germany and the German Association of the Automotive Industry (VDA), January* 26, 2016, disponibile all'indirizzo: https://www.lda.bayern.de/media/dsk_joint_statement_vda.pdf;

[30] CNIL, *Compliance package. Connected vehicles and personal data, October 2017 Edition*.

[31] L. MIGLIETTI, *Profili storico-comparativi del diritto alla privacy*, in *diritticomparati.it*, 2014, disponibile all'indirizzo: https://www.diritticomparati.it/profili-storico-comparativi-del-diritto-alla-privacy/

contrapponevano i coltivatori americani timorosi che i dati raccolti dai sensori incorporati nei macchinari agricoli intelligenti potessero finire nelle mani sbagliate e causare esternalità negative nei loro confronti, sia in termini di perdita di potere contrattuale che di scoperta di eventuali illeciti che avrebbero portato a provvedimenti sanzionatori.[32] Malgrado siano emerse anche considerazioni attinenti alla *privacy* dei soggetti coinvolti, la concezione restrittiva di dato personale diffusa negli Stati uniti ha indirizzato la discussione non tanto sulla qualificazione delle informazioni raccolte, quanto piuttosto sulle questioni inerenti alla loro proprietà ed al loro utilizzo.[33] La stessa soluzione di compromesso che è stata raggiunta tra le associazioni di rappresentanza degli agricoltori ed i principali *agricultural technology provider*, consistente in un accordo che stabilisce una serie di principi guida che definiscono in maniera più puntuale la cornice di trattamento di questo tipo di informazioni, si focalizza sui profili commerciali dell'analisi dei dati agricoli, senza soffermarsi sui rischi di identificabilità derivanti dal trattamento di tali dati.[34]

In considerazione di tale paragone, risulta del tutto evidente l'influenza che differenti esperienze e criteri possono avere nei confronti del circuito della decisione algoritmica: se l'impostazione restrittiva statunitense è in grado di fornire all'analisi algoritmica un paniere di *input* molto più ampio da cui attingere per il *training*, l'approccio "dato-personale-centrico" europeo, nell'intento di tutelare maggiormente i diritti fondamentali della persona fisica, rischia di frapporre nuovi ostacoli allo sviluppo dei sistemi algoritmici nel vecchio continente.[35]

3.2 In cerca di una risposta per i nuovi quesiti posti dai Big Data

Nell'ambito dell'analisi dei sistemi algoritmici, un settore che merita particolare considerazione è quello dei Big Data, oggetto di esame del secondo documento in discussione. Il cambio di rotta realizzatosi con l'arrivo dell'era dei Big Data risiede tutto nel passaggio dal precedente metodo logico-deduttivo, dove l'algoritmo trae delle conclusioni sulla base delle istruzioni che gli sono state fornite esclusivamente dall'essere umano, ad un nuovo approccio metodologico in cui attraverso l'osservazione dei dati è la macchina stessa ad apprendere e, in seguito, predire con una alta probabilità statistica il verificarsi di

[32] OECD, *Enhancing Access to and Sharing of Data: Reconciling Risks and Benefits for Data Re-use across Societies*, disponibile all'indirizzo: https://www.oecd.org/sti/enhancing-access-to-and-sharing-of-data-276aaca8-en.htm

[33] M. SYKUTA, *Big Data in Agriculture: Property Rights, Privacy and Competition in Ag Data Services*, in *International Food and Agribusiness Management Review*, Special Issue, Vol. 19 Issue A, 2016, 57 ss.

[34] *Ag Data's Core Principles: The Privacy and Security Principles for Farm Data, Ag Data Transparent*, 2016, disponibile all'indirizzo: http://www.agdatatransparent.com/principles/

[35] P. SCHWARTZ, D. SOLOVE, *The PII Problem*, cit.

determinati scenari.[36] È stato inevitabile, pertanto, che il nuovo paradigma *data driven* si ripercuotesse sul valore che hanno i dati nella società odierna: anche dati prima privi di particolare valenza, ad esempio perché meri *by-product* di altri procedimenti di calcolo, ora acquistano un significato a sé stante, giacché è nell'individuazione di *pattern* che sfuggono alle capacità computazionali finora conosciute che si sostanzia la vera anima dell'analisi dei Big Data. Questa onda nuova ha riguardato tutti i dati che possono essere analizzati dagli algoritmi, dunque anche – *rectius*, soprattutto – quelli non personali, i quali contribuiscono in quota maggioritaria a comporre quello sterminato cluster di informazioni da cui la *data analytics* ricava la sua materia prima. Tuttavia, dal punto di vista della qualificazione normativa, l'accorpamento di moli di informazioni così vaste e le possibilità di re-identificazione assicurate dai nuovi strumenti tecnologici utilizzati per la loro gestione hanno acuito le problematiche di accertamento della natura dei dati, la quale si mostra sempre più mutevole e dinamica.[37]

Un preciso riconoscimento del ruolo ricoperto dai dati non personali nell'ambito di questa rivoluzione è ravvisabile nella "Indagine conoscitiva sui Big Data" del 2018, condotta congiuntamente da tre autorità italiane: il Garante per la protezione dei dati personali, l'Autorità per le Garanzie nelle Comunicazioni (AGCOM) e l'Autorità Garante della Concorrenza e del Mercato (AGCM).[38] Già nella parte introduttiva, lo studio, volto a definire il fenomeno dei Big Data ed a misurare il suo impatto sull'ecosistema digitale, sottolinea la centralità della distinzione tra le due categorie di dati nell'identificazione del regime giuridico applicabile.[39]

La visione che traspare dal rapporto nasce proprio dalle considerazioni svolte in merito allo sviluppo dell'analisi algoritmica. Secondo i garanti, il processo di estrazione di conoscenza reso possibile dalla mole di informazioni oggi a disposizione avviene tramite un approccio nuovo "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".[40] Dunque, se in un primo momento i dati avevano acquisito un significato normativo in quanto collegati alla tutela di un diritto fondamentale,[41] ora assumono un valore nuovo ed

[36] A. SIMONCINI, S. SUWEIS, *Il cambio di paradigma nell'intelligenza artificiale e il suo impatto sul diritto costituzionale*, in *Rivista di filosofia del diritto*, Fascicolo 1, 2019, 87 ss.

[37] Consiglio d'Europa, *Guidelines on the protection of individuals with regard to the processing of personal data in a world of Big Data*, 2017.

[38] AGCM, AGCOM, Garante per la protezione dei dati personali, *Indagine conoscitiva sui big data*, 2018, disponibile a: https://www.agcm.it/dotcmsdoc/allegati-news/IC_Big%20data_imp.pdf

[39] *Ibid*., 8-10.

[40] *Ibid*., 17.

[41] In particolare, i dati personali avevano ottenuto maggiore significato normativo a seguito del riconoscimento dello stesso rango dei trattai istitutivi dell'Unione europea della Carta dei diritti

ulteriore che non concerne direttamente la dimensione della tutela personale, ma che attiene allo sviluppo di algoritmi i cui *output* possono comunque produrre un effetto considerevole sulla sfera giuridica dell'individuo. Alla luce di tali osservazioni, i garanti dedicano i successivi paragrafi all'individuazione delle principali sfide che questo cambio di paradigma pone e, stante l'affanno nel trovare soluzioni percorribili *de jure condito*, si spingono a proporre alcune alternative agli schemi classici finora utilizzati, troppo spesso non adeguati a governare una realtà così complessa come quella dei Big Data.

Innanzitutto, il documento affronta il tema dell'anonimizzazione come possibile soluzione di compromesso tra la tutela della persona fisica e l'esigenza di enti pubblici e privati di avere accesso ai dati per analizzarli ed orientare le proprie strategie. In particolare, le autorità esaminano le argomentazioni di quella parte della dottrina e dell'industria che erge l'anonimizzazione a via maestra per garantire l'equilibrio tra interessi contrapposti, sulla base della logica secondo cui gli operatori del settore digitale, specialmente nell'ambito del *marketing*, non sono interessati ad identificare la persona alla quale originariamente i dati si riferivano, ma, diversamente, a creare dei "tipi-ideali" di individuo, all'interno dei quali possano rientrare soggetti con preferenze e usi simili.[42] Tuttavia, le stesse autorità manifestano più di una perplessità al riguardo dal punto di vista giuridico, giacché descrivere le attività di profilazione eseguite dal titolare del trattamento come via d'uscita dall'osservanza degli obblighi giuridici posti a tutela delle persone fisiche si rivelerebbe contrario alla *ratio* che ispira l'intero Regolamento europeo sulla protezione dei dati personali.[43] Pertanto, la costruzione di modelli ideali elaborati per mezzo di procedure di anonimizzazione non implica una automatica fuoriuscita dall'ambito della disciplina sulla protezione dei dati personali, tanto più se si considera che proprio l'utilizzo intensivo degli algoritmi ha pregiudicato ulteriormente l'irreversibilità dei processi di anonimizzazione.[44] In ragione dei rischi di insuccesso connessi alle procedure di de-identificazione, le autorità sembrano riconoscere, tra le righe, che la mancanza di un sistema di regole omogeneo che definisca in maniera più chiara le misure tecniche e giuridiche, anche settoriali, per pervenire ad una effettiva anonimizzazione abbia un costo oneroso per il nostro ordinamento. Un passo in avanti in questa direzione non andrebbe solamente a beneficio della protezione dei dati personali, ma, per riprendere le parole delle

fondamentali dell'Unione europea, e dell'introduzione dell'art. 16 TFUE. A tale riguardo, si veda: G. GONZALEZ FUSTER, *The emergence of personal data protection as a fundamental right of the EU*, Springer, Cham, Law, Governance and Technology Series, 2014.

[42] AGCM, AGCOM, Garante, *Indagine conoscitiva sui big data*, cit., 23-25.

[43] R. HU ET AL., *Bridging Policy, Regulation and Practice?*, cit.

[44] P. SCHWARTZ, D. SOLOVE, *The PII Problem*, cit.

autorità, si rivelerebbe essenziale "anche nell'ottica di coerenza con la strategia nazionale di sicurezza cibernetica".[45]

Un altro aspetto su cui si sofferma l'indagine dei garanti concerne la parziale dissonanza presente tra alcuni dei cardini della disciplina europea, il cui baricentro è pesantemente spostato verso il lato personale dei dati e verso la fase anteriore del trattamento, e le dinamiche dei Big Data, dove tanto la natura delle informazioni quanto le finalità del trattamento sono difficilmente identificabili *ex ante*.[46] Sfortunatamente, una volta individuato il problema, le autorità non riescono a tracciare una strada che permetta di superare in modo soddisfacente le problematicità introdotte dalla duplice qualificazione normativa delle informazioni. Se in un primo momento, si ribadisce la necessità di identificare la natura del dato in via preliminare, ossia prima dell'inserimento nell'algoritmo, per conoscere il regime giuridico rilevante, successivamente, quando si tenta di definire le modalità attraverso cui procedere ad una chiara distinzione, il documento si limita ad un generico richiamo alle migliori prassi in materia di anonimizzazione, oppure a provvedimenti o orientamenti – non meglio precisati – delle autorità nazionali ed europee, di fatto soprassedendo alla soluzione di una questione di enorme impellenza.[47]

Invero, a riconferma delle difficoltà insite nella ricerca di una soluzione adeguata, in un passaggio dell'indagine che riprende quanto dichiarato nell'Interim Report elaborato dall'AGCOM nella fase preparatoria,[48] le autorità arrivano ad ipotizzare un aggiornamento del quadro regolamentare tramite una normativa in materia di dati che non si basi sulla attuale classificazione binaria, ma che, al fine di ovviare alle complicazioni inerenti alle operazioni di distinzione, miri "a proteggere i dati da un punto di vista generale".[49] Questa idea di superare la divisione tra dato personale e non personale, non dissimile da quanto prospettato anche in altre sedi del contesto europeo,[50] oltre che mostrare l'insofferenza degli stessi organi regolatori di fronte ad una distinzione spesso inafferrabile, potrebbe costituire la base per un eventuale sviluppo di un filone di pensiero destinato ad attirare maggiori sostenitori negli anni a venire; anche se, per il momento, appare improbabile che il legislatore torni sui propri passi per rimettere mano a testi normativi piuttosto recenti.

[45] AGCM, AGCOM, Garante, *Indagine conoscitiva sui big data*, cit., 117.

[46] *Ibid*., 25-26.

[47] *Ibid*., 8-10.

[48] Big data Interim report nell'ambito dell'indagine conoscitiva di cui alla delibera n. 217/17/CONS, AGCOM, 2018.

[49] AGCM, AGCOM, Garante, *Indagine conoscitiva sui big data*, cit., 45-46.

[50] Parere del Comitato economico e sociale europeo (CESE) sulla "Comunicazione della Commissione al Parlamento europeo e al Consiglio - Linee guida sul regolamento relativo a un quadro applicabile alla libera circolazione dei dati non personali nell'Unione europea" adottato il 25 settembre 2019.

3.3 Prime osservazioni sull'opportunità di regolare la decisione algoritmica in Francia

Fra i documenti espressamente dedicati allo studio degli effetti che la decisione algoritmica può produrre sulla sfera giuridica individuale, sono stati selezionati quelli in cui le autorità amministrative indipendenti dimostrano maggiore consapevolezza dell'incidenza che la fattispecie dei dati non personali può avere in questo ambito. Particolarmente interessante in tal senso è il rapporto elaborato dalla CNIL nel 2017, dove, a seguito di una consultazione pubblica, l'autorità francese tenta di rispondere agli interrogativi sollevati dalla diffusione dell'utilizzo degli algoritmi a fini decisionali o consultivi ed elabora alcune raccomandazioni etico-giuridiche aventi lo scopo di garantire la salvaguardia dei diritti e delle libertà degli individui.[51]

In esito all'analisi relativa all'inquadramento giuridico, fra le numerose normative che, direttamente o indirettamente, condizionano l'utilizzo degli algoritmi, il documento si pone sulla stessa linea della dottrina maggioritaria e si focalizza sulla disciplina che ha l'impatto più significativo, quella relativa al trattamento dei dati personali, sia nazionale che europea. Effettivamente, non potrebbe essere altrimenti atteso che il regolamento europeo e le normative nazionali che ne sono derivate contengono disposizioni di principio appositamente introdotte al fine di disciplinare le decisioni basate sul trattamento automatizzato dei dati a carattere personale.[52] Nello specifico, fra le norme centrali che meritano di essere richiamate, vi sono, dal punto di vista generale, i principi che reggono il trattamento delle informazioni personali e che si applicano a qualsiasi tipo di trattamento, dunque, anche a quello algoritmico,[53] mentre, scendendo nel dettaglio, spiccano il diritto di essere informati circa l'esistenza di un processo decisionale automatizzato,[54] ed il divieto generale – tra l'altro, derogato da numerose eccezioni – di adottare una decisione basata unicamente sul trattamento automatizzato che incida significativamente su una persona fisica.[55]

D'altro canto, malgrado le guarentigie predisposte dal GDPR siano comunque da apprezzare in qualità di primo baluardo a difesa della sfera individuale dell'essere umano, limitare l'analisi alla sola disciplina relativa ai dati a carattere personale restituisce una visione alquanto parziale delle modalità attraverso cui l'impianto continentale cerca di regolare il vasto campo della decisione

[51] CNIL, *How can humans keep the upper hand? The ethical matters raised by algorithms and artificial intelligence*, 2017.

[52] A. SIMONCINI, *L'algoritmo incostituzionale*, cit.

[53] Artt. da 5 a 11, GDPR.

[54] Art. 15, par. 1, lett. h), GDPR.

[55] Art. 22, GDPR.

algoritmica e non permette di individuare i punti deboli del sistema.[56] Consapevole di ciò, nel prosieguo della disamina l'autorità francese localizza alcune dimensioni tipiche della decisione automatizzata che non sono state prese in debita considerazione dal legislatore. In primo luogo, il garante si concentra sull'assenza di previsioni generali applicabili a quegli algoritmi che elaborano solamente dati non personali e che possono ugualmente incidere sui diritti degli individui. A tale proposito, entra in gioco il concetto di scala di applicazione della decisione frutto di un trattamento automatizzato: la stessa scelta presa in ragione delle risultanze fornite da un algoritmo attraverso l'analisi di dati non personali non genera, di regola, particolari conseguenze se viene riservata ad un ambito piuttosto circoscritto, ma può avere ripercussioni travolgenti nel caso in cui venisse applicata su scala più ampia, ad esempio, all'intera popolazione di un Paese.[57] D'altronde, ricadute altrettanto rilevanti di questa tipologia di decisioni possono aversi anche nelle ipotesi in cui il destinatario sia un singolo individuo. Basti pensare, a tale riguardo, alle decisioni prese dagli algoritmi incorporati nelle automobili a guida autonoma basate sulle condizioni della strada, sul meteo e sui movimenti degli altri veicoli o al calcolo delle quantità di prodotti da utilizzare nell'agricoltura di precisione: un errore di valutazione da parte della macchina porterebbe ad effetti dannosi potenzialmente molto gravi.[58] A questa prima dimensione se ne collega una seconda, anch'essa priva di adeguata copertura giuridica, che riguarda i casi in cui i destinatari delle decisioni algoritmiche non siano i singoli individui, ma intere comunità. Difatti, benché sia stato chiarito dalla dottrina che il GDPR trova applicazione in tutte quelle ipotesi in cui, benché basata su una profilazione di gruppo, la decisione sia diretta ad incidere in maniera significativa sulla sfera giuridica di un individuo,[59] secondo l'autorità permangono alcuni dubbi nel caso in cui gli effetti si producano su di un insieme di soggetti complessivamente considerato,

[56] B. SCHERMER, *The limits of privacy in automated profiling and data mining*, in *Computer law & security review* 27, 2011, 45 ss.

[57] L'autorità riprende un esempio elaborato da una ricercatrice americana, la quale ipotizza la progettazione di un algoritmo che, analizzando dati a carattere non personale, sia in grado di indicare il regime alimentare più salutare per i bambini di una famiglia. Tuttavia – prosegue l'autrice – nel caso in cui la medesima dieta risultante dal calcolo algoritmico fosse applicata, anziché a singole famiglie, a tutti gli studenti delle mense scolastiche degli Stati Uniti, si genererebbero effetti economici e sociali dirompenti sulla produzione, sui costi e sull'occupazione del settore agro-alimentare. A tal proposito, si veda: C. O'NEIL, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, Crown, 2016.

[58] M. BRKAN, G. BONNET, *Legal and Technical Feasibility of the GDPR's Quest for Explanation of Algorithmic Decisions: of Black Boxes, White Boxes and Fata Morganas*, in *European Journal of Risk Regulation*, Vol. 11, Issue 1, 2020, 18 ss.

[59] G. NOTO LA DIEGA, *Against the Dehumanisation of Decision-Making – Algorithmic Decisions at the Crossroads of Intellectual Property, Data Protection, and Freedom of Information*, in JIPITEC, 2018, 3 ss.

senza che la dimensione del singolo assuma rilievo e senza, dunque, che le modalità di esercizio dei diritti che la normativa europea garantisce alla persona fisica siano sufficientemente chiare.[60]

Il merito di questo studio della CNIL risiede proprio nell'essere riuscito a cogliere, sebbene in maniera ancora piuttosto acerba, uno dei punti fondamentali della discussione inerente alla regolazione della decisione algoritmica. Il mancato apprezzamento da parte del legislatore delle lacune descritte pone l'individuo e la collettività in una posizione di netto svantaggio rispetto al soggetto decisore, sia esso un organo pubblico oppure privato. Trascurare i sistemi algoritmici che si nutrono di dati a carattere non personale significa impedire a coloro che ne subiscono gli effetti di esercitare alcuni dei diritti basilari dell'epoca della digitalizzazione, su tutti il diritto di accesso e il diritto alla cosiddetta "explainability", di fatto contravvenendo sia al principio di trasparenza che a quello di responsabilità.[61] Gli altri strumenti giuridici cui potrebbe ricorrere l'individuo leso dalla decisione algoritmica, fra i quali figurano il diritto di accesso del diritto amministrativo, le eccezioni in materia di diritto d'autore,[62] o persino la *privacy law* – intesa quale branca distinta rispetto al diritto alla protezione dei dati personali, poiché diretta a tutelare la vita privata come libertà negativa di impedire interferenze esterne –[63] mostrano comunque grandi limiti sia perché basati su particolari presupposti normativi non sempre sussistenti in questi casi, sia perché non sono stati originariamente concepiti come meccanismi di difesa dal trattamento automatizzato di informazioni.

Alla luce di tali considerazioni, l'autorità francese conclude con una serie di raccomandazioni che spaziano dall'opportunità di introdurre nuovi obblighi o divieti in capo ai *designer* ed agli utilizzatori degli algoritmi, all'implementazione di meccanismi di *auditing*, sino ad escluderne l'impiego in determinati settori in ragione della loro sensibilità. Tuttavia, la stessa CNIL si dimostra favorevole a colmare le lacune della disciplina attualmente in vigore per mezzo di un approccio misto che non si esaurisca esclusivamente nell'introduzione di un *corpus* di disposizioni a carattere vincolante, ma che sia capace di abbracciare anche soluzioni volontarie. L'obiettivo ultimo del metodo *bottom-up* sponsorizzato dal garante è quello di superare la cosiddetta "*silo mentality*" attraverso un percorso di formazione etica e di coinvolgimento di tutti i soggetti che operano nella catena di costruzione dell'algoritmo: chi progetta il

[60] CNIL, *How can humans keep the upper hand?,* cit., 46.

[61] M. BRKAN, G. BONNET, *Legal and Technical Feasibility of the GDPR's Quest for Explanation of Algorithmic Decisions*, cit.

[62] G. NOTO LA DIEGA, *Against the Dehumanisation of Decision-Making*, cit.

[63] M. ALMADA, M. DYMITRUK, *Privacy and Data Protection Constraints to Automated Decision-Making in the Judiciary*,in SSRN Electronic Journal, 2020.

codice, chi lo utilizza per decidere, fino ad arrivare a chi è destinato a subire gli effetti della decisione.[64]

3.4 L'approccio normativo nell'opinione della *Data Ethics Commission* tedesca

Un ulteriore passo in avanti, indice di una concezione più matura e consapevole dell'universo delle decisioni algoritmiche, è ravvisabile nell'ultimo documento esaminato. Da un punto di vista prettamente formale, esso non costituisce opera esclusiva di un'autorità amministrativa indipendente, ma è frutto del lavoro congiunto di esponenti provenienti dai settori accademico, amministrativo ed industriale, fra i quali figurano membri del garante per la protezione dei dati tedesco, sia federale che del Land dello Schleswig-Holstein.[65] Il gruppo, denominato *Data Ethics Commission* (*Datenethikkommission*), è stato istituito dal governo tedesco nel 2018 allo scopo di elaborare alcune proposte in riferimento ai principali quesiti etico-giuridici afferenti a tre macroaree: le decisioni algoritmiche, l'intelligenza artificiale ed i dati. Ciononostante, la Commissione ha ridotto le aree di intervento a due, sulla base dell'assunto secondo cui il settore dell'intelligenza artificiale altro non è che una delle diramazioni di cui si compone lo sterminato ambito dei sistemi algoritmici. Dunque, partendo da questa premessa, l'opinione si struttura in due parti: la prima dedicata alla fase degli *input*, quindi dei dati; la seconda concernente il sistema algoritmico in generale, ricomprendendo dunque sia la fase di calcolo che i successivi effetti.[66]

La Commissione dimostra piena contezza delle molteplici sfaccettature in cui si articola l'universo dei dati ed enuncia sin dal principio la centralità della *summa divisio* tra dati personali e non personali per la definizione del regime giuridico rilevante. Tuttavia, a differenza di quanto avviene negli altri documenti riportati, in questo caso le osservazioni degli autori non si esauriscono nell'enunciazione delle carenze dell'ordinamento, ma arrivano a proporre alcune regole concrete specificamente dedicate al versante dei dati non personali. A tale scopo, la Commissione formula alcuni suggerimenti interessanti come l'introduzione di una presunzione legale di non personalità del dato quando le

[64] CNIL, *How can humans keep the upper hand?*, cit., 61. In particolare, la "*silo mentality*" è definita come una delle tre principali sfide della società digitale, assieme all'imprevedibilità degli algoritmi ed alla eccessiva fiducia nelle macchine, e viene così descritta: "the silo mentality affect[s] the organisation of algorithmic chains, which leads to action being carried out in isolation, indifference to the overall impacts of the algorithmic system and diminishing accountability".

[65] Opinion of the Data Ethics Commission, 2019, disponibile all'indirizzo: https://www.bmjv.de/SharedDocs/Downloads/DE/Themen/Fokusthemen/Gutachten_DEK_EN.pdf?__blob=publicationFile&v=2#:~:text=In%20the%20opinion%20of%20the%20Data%20Ethics%20Commission%20such%20contributions,the%20part%20of%20other%20parties.

[66] *Ibid.*, 13.

procedure di anonimizzazione sono state condotte in conformità a standard riconosciuti, l'inserimento di apposite sanzioni penali per coloro che tentano la re-identificazione, la previsione espressa di "*data-specific rights of co-determination and participation*" a beneficio di coloro che hanno contribuito alla generazione dell'informazione – dai quali, tuttavia, viene espressamente escluso il diritto di proprietà – e, non da ultimo, il riconoscimento di un diritto alla protezione dei dati anche alle persone giuridiche, nell'ottica di tutela della sovranità digitale delle imprese del nostro continente.[67]

Il punto centrale del rapporto tra dati non personali e decisione algoritmica colto dalla Commissione in questa prima parte concerne la necessità di superare quell'approccio "libertario" in ragione del quale si ritiene possibile qualsiasi tipo di utilizzo dei dati una volta al di fuori dell'ambito di applicazione del GDPR. Tale impostazione, che sembra permeare tanto la prassi degli operatori di settore quanto il panorama legislativo, si rivela del tutto inadeguata in un momento storico in cui agli algoritmi viene garantito uno spazio sempre maggiore. Con particolare riguardo alle modalità di utilizzo dei dati di *input*, se è vero che i sistemi algoritmici riproducono i pregiudizi già presenti nell'insieme di dati di partenza, allora si rivela opportuno, se non addirittura necessario, dettare alcune regole in merito, fra l'altro, alla provenienza di tali dati, prestando particolare attenzione al fatto che l'impiego di dataset generati nel contesto di un determinato ambiente rischiano di produrre risultati inesatti e persino deleteri quando vengono utilizzati come risorsa per prendere decisioni in ambiti totalmente differenti.[68] Dunque, nella prospettiva della regolazione della decisione algoritmica, la struttura predisposta dal GDPR si dimostra inadeguata a prevenire ogni effetto discriminatorio possibile, poiché si focalizza principalmente sull'imposizione al titolare del trattamento dell'obbligo di fare ricorso a procedure sicure che rispettino i principi di esattezza e di minimizzazione, mentre non dedica opportuna attenzione a quali tipi di dati dovrebbero essere inclusi o esclusi al preciso scopo di prevenire tale discriminazione.[69] Difatti, come è stato correttamente osservato, sono proprio le proibizioni imposte dal Regolamento all'utilizzo di alcuni dati personali – su tutti, il divieto di utilizzo di categorie particolari di dati di cui all'art. 9 – che a volte rischiano di

[67] *Ibid*., 17, 79 ss. In tal caso, gli autori ipotizzano una sorta di parallelismo tra la disciplina dei dati personali e quella dei dati non personali, in particolare: "Digital self-determination in the data society also includes the self-determined economic exploitation of one's own data, and it includes self-determined management of non-personal data, such as non-personal data generated by one's own devices. The Data Ethics Commission takes the view that, in principle, a right to digital self-determination in the data society also applies to companies and legal entities and – at least to some extent – to groups of persons (collectives)".

[68] P. DE LAAT, *Algorithmic Decision-Making Based on Machine Learning from Big Data: Can Transparency Restore Accountability?*, in *Philos. Technol*. 31, 2018, 525 ss.

[69] G. NOTO LA DIEGA, *Against the Dehumanisation of Decision-Making*, cit.

pregiudicare irrimediabilmente l'accuratezza e l'imparzialità della decisione algoritmica.[70]

Passando alla seconda parte dell'opinione, la Commissione, sulla base della premessa secondo cui gli algoritmi che processano solamente dati a carattere non personale possono produrre effetti altrettanto significativi nei confronti degli individui, si sofferma sulla necessità di considerare separatamente l'ambito del trattamento dei dati, da un lato, e quello della decisione algoritmica, dall'altro.[71] Di conseguenza, l'approccio improntato all'osservanza del fenomeno non solo nella prospettiva *ex ante* ma anche *ex post* suggerisce di procedere con una impostazione binaria dal punto di vista normativo, aggiungendo alla disciplina relativa al trattamento dei dati un nuovo regolamento europeo specificamente dedicato alla disciplina degli algoritmi. Il "*Regulation on Algorithmic Systems*" proposto dalla Commissione, consistente in un insieme di principi base applicabili orizzontalmente a tutte le normative di settore supplementari riguardanti decisioni automatizzate, poggia su un "*risk-adapted regulatory approach*" che, similmente a quanto fa il GDPR in riferimento ai dati personali,[72] prenda in considerazione qualsiasi rischio collegato all'utilizzo di sistemi algoritmici, comprendendo anche quelli derivanti dal trattamento di dati non personali. In termini normativi, la collocazione del fattore di rischio in uno dei cinque livelli previsti dalla proposta – sulla base della probabilità del verificarsi di un evento dannoso e della relativa gravità del danno – determina l'intensità delle limitazioni da applicare, le quali possono andare da restrizioni minime, fino alla proibizione, totale o parziale, della decisione algoritmica.[73] Oltretutto, l'elemento del rischio non esaurisce la sua rilevanza con riguardo al ventaglio di utilizzi giuridicamente consentiti, poiché è capace di incidere anche sul tipo di strumento prescelto per la regolamentazione: una probabilità particolarmente bassa che si verifichi un evento dannoso apre la strada a misure di *self-regulation* che mitigherebbero l'atteggiamento maggiormente prescrittivo dimostrato dalla Commissione tedesca.[74]

In sostanza, in base al presupposto secondo cui anche trattamenti di dati che non abbiano nessuna connessione con una persona fisica specifica possono produrre effetti significativi sulla sfera giuridica individuale, la Commissione si pone l'obiettivo di rivedere l'impianto regolamentare in vigore, attualmente costituito solo da un primo livello di protezione, il Regolamento sulla protezione dei dati personali, il quale, una volta superato, lascia gli interessati inermi

[70] B. SCHERMER, *The limits of privacy in automated profiling and data mining*, cit.

[71] Opinion of the Data Ethics Commission, cit., 157 ss.

[72] S. CALZOLAIO, *Protezione dei dati personali*, in R. BIFULCO, A. CELOTTO, M. OLIVETTI (a cura di) *Digesto delle Discipline Pubblicistiche*, Utet giuridica, 2017, 594 ss.

[73] Opinion of the Data Ethics Commission, cit., 173 ss.

[74] *Ibid*., 201-204.

di fronte alle conseguenze della decisione algoritmica. Malgrado l'assenza di violazioni al GDPR – o perché le sue disposizioni sono state rispettate o perché si è al di fuori del suo campo di applicazione – possono comunque aversi ripercussioni pregiudizievoli nei confronti delle quali i destinatari hanno ben pochi strumenti a disposizione per opporsi.[75] In siffatti casi, alcuni dei principi che dovrebbero, comunque, presidiare l'attuazione della decisione algoritmica rischiano di non reggere l'impatto con la potenza della tecnologia.[76] Il principio di non discriminazione non è esente da potenziali violazioni, giacché gli effetti discriminatori potrebbero accidentalmente derivare, oltre che dai dati di *input* utilizzati, anche dal modo in cui l'algoritmo ha imparato ad associarli nella fase di *learning*, secondo modalità spesso poco chiare alla mente umana.[77] Le medesime considerazioni valgono per il principio di trasparenza, i cui corollari relativi ai diritti di essere informati, di accesso e di ottenere una spiegazione cadono in un vuoto normativo quando si tratta di decisioni basate su dati puramente non personali.[78] In aggiunta, il pericolo di ingiusta lesione delle prerogative individuali viene amplificato enormemente dall'assenza di una specifica autorità di controllo deputata a monitorare il corretto utilizzo dei sistemi algoritmici e ad individuare eventuali responsabilità in caso di evento dannoso.[79] Perciò, la regolazione dei sistemi algoritmici dovrà riguardare sia il profilo relativo alle autorità di controllo, attraverso un allargamento delle competenze di quelle già esistenti o, persino, dando vita ad organi del tutto nuovi,[80] sia il problema della allocazione delle responsabilità per le imprese che operano nell'ambito delle tecnologie digitali.[81]

In definitiva, un'eventuale riforma dell'ordinamento europeo non dovrebbe limitarsi solamente ad una modifica della normativa in tema di trattamento dei dati atta a regolamentare anche l'utilizzo di quelli non personali (*data perspective*) ma dovrebbe incorporare, in chiave complementare, una disciplina incentrata sulla fase relativa agli effetti delle decisioni prodotte dall'algoritmo

[75] B. SCHERMER, *The limits of privacy in automated profiling and data mining*, cit.

[76] A. SIMONCINI, S. SUWEIS, *Il cambio di paradigma nell'intelligenza artificiale*, cit.

[77] B. SCHERMER, *The limits of privacy in automated profiling and data mining*, cit.

[78] M. BRKAN, G. BONNET, *Legal and Technical Feasibility of the GDPR's Quest for Explanation of Algorithmic Decisions*., cit.

[79] B. SCHERMER, *The limits of privacy in automated profiling and data mining*, cit.

[80] Opinion of the Data Ethics Commission, cit., 198 ss.

[81] *Ibid*., 70-71. La Commissione fa riferimento alla "Corporate Digital Responsibility" (CDR), la quale viene descritta come "[…] the idea that companies, as manufacturers and operators of digital technologies, should each assume their own responsibility for the consequences of digitalisation. Like corporate social responsibility (CSR), CDR falls under the broader umbrella of corporate responsibility; in this case, the focus is on voluntary corporate activities in the digital sphere which go beyond what is currently prescribed by law, and which actively shape the digital world to the benefit of society in general, and of customers and employees in particular".

(*algorithms perspective*).[82] In questo senso, alla Commissione va riconosciuto il merito di aver identificato brillantemente queste due dimensioni che, essendo distinte e, al contempo, strettamente dipendenti, devono essere trattate in maniera separata, ma tenendo conto della loro influenza reciproca. Da un lato, la (parziale) assenza di regolamentazione per i dati a carattere non personale non consente di tutelare tutti i soggetti coinvolti nella catena di *governance* delle informazioni, specie le persone giuridiche, e ciò, nell'era della globalizzazione e della digitalizzazione, rischia di pregiudicare la sovranità digitale degli Stati membri e dell'Europa intera. Dall'altro lato, se l'ordinamento europeo vuole offrire una protezione integrale agli interessati non può focalizzarsi esclusivamente sulla fase anteriore di trattamento dei dati, ma deve predisporre un meccanismo di difesa nei confronti degli *output* dell'algoritmo, che permetta dunque di intervenire anche nei casi in cui, malgrado la disciplina sul trattamento dei dati sia stata rispettata, si siano prodotti effetti indesiderati e discriminatori.

4. Osservazioni conclusive

La multidimensionalità e l'interdisciplinarietà della decisione algoritmica sono fattori che contribuiscono a complicare i tentativi di inquadramento giuridico del fenomeno. Se il lavoro della dottrina si è finora focalizzato principalmente sulla disciplina relativa al trattamento dei dati a carattere personale in ragione di alcune delle più importanti disposizioni inserite nel GDPR con specifico riguardo alla fattispecie in discussione, osservare i sistemi algoritmici attraverso la lente della fattispecie dei dati non personali rappresenta un'interessante occasione per riflettere sulle lacune dell'impianto europeo e, al contempo, permette di avere contezza di quanto sia necessario perfezionare l'ordinamento a doppia velocità predisposto dal legislatore in materia di dati. Più precisamente, il presente contributo è stato redatto con l'intento di approfondire la prospettiva delle autorità amministrative indipendenti europee tramite lo studio di quei documenti che sono riusciti a mettere in risalto in maniera più nitida la stretta relazione che lega i dati non personali alla decisione algoritmica. Il percorso seguito mostra in maniera piuttosto evidente il passaggio dalla ricerca di soluzioni all'interno del diritto positivo alla proposta di nuovi strumenti normativi maggiormente corrispondenti alle nuove sfide poste dallo sviluppo tecnologico degli algoritmi. Se il concetto di identificabilità si è rivelato un mezzo utile in tutte quelle situazioni limite in cui la qualifica dell'informazione appare incerta, contemporaneamente, esso si caratterizza per alcuni effetti pregiudizievoli che sul lungo periodo rischiano di minare quelle politiche economiche dell'Unione europea volte a recuperare parte del potere perso nei confronti delle altre potenze mondiali, orientali ed occidentali. Inoltre, tanto in

[82] *Ibid.*, 77.

riferimento al versante dei Big Data quanto a quello dei sistemi algoritmici in generale, le autorità amministrative indipendenti si stanno allontanando progressivamente dalla *law in the books*, spesso poco consona alle dinamiche della digitalizzazione, per proiettarsi verso la *law in action*, con la proposizione di soluzioni capaci di riempire i vuoti esistenti tra disciplina giuridica vigente e realtà dei fatti. Le nuove regole riguardanti i procedimenti di anonimizzazione, la riforma relativa al trattamento dei dati non personali e la previsione di un regolamento *ad hoc* per i sistemi algoritmici sono tutti indicatori di una certa insofferenza patita dagli organi di controllo innanzi ad un fenomeno che muta così velocemente da rendere assai arduo assolvere ai doveri di monitoraggio tramite le regole statiche al momento applicabili.

Invero, è opportuno sottolineare a tal proposito come la proposta di regolamento europeo sull'intelligenza artificiale recentemente approvata dalla Commissione europea sembri muoversi in questa direzione. Sebbene il testo definitivo non vedrà la luce prima di alcuni anni, si percepisce aria di svolta in ambito continentale poiché la proposta, dedicando specifica attenzione alle fasi di progettazione e successiva immissione nel mercato dei sistemi di intelligenza artificiale, prescinde, per quanto possibile, dalla separazione tra dati personali e non personali. Difatti, le disposizioni che la compongono rappresentano il primo vero tentativo di risposta ad alcuni degli interrogativi sollevati nel corso del presente lavoro; segnatamente, la limitazione degli effetti lesivi dei diritti fondamentali derivanti dall'utilizzo dell'intelligenza artificiale, l'obbligo di assicurare una alta qualità dei dataset utilizzati per il *training* e l'attribuzione di peculiari compiti di controllo ad autorità indipendenti.[83]

Facendo riserva di svolgere le opportune considerazioni quando la proposta terminerà il suo iter legislativo, è possibile sin da ora affermare che sussiste una certa convergenza di opinioni sulla necessità di agire in tema di intelligenza artificiale, anche se, peraltro, permane qualche dubbio in riferimento alla forma ed al contenuto che dovrebbe assumere tale intervento. Il dibattito fra i vari *stakeholder* sul rapporto tra lo spazio ricoperto dalla normazione per mezzo di una fonte di rango primario e quello lasciato a meccanismi di co-regolamentazione o auto-regolamentazione non è stato ancora totalmente superato. Pertanto, non sarà facile trovare un "equilibrio europeo", dove la spinta verso uno sviluppo tecnologico più celere non venga ottenuta a costo di sacrificare sia i valori fondamentali su cui è stato costruito l'ordinamento continentale sia la sovranità dei suoi Stati membri, mai così a rischio come oggi.

[83] Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts, Brussels, COM(2021) 206 final.

Altrettanto delicate sono le questioni che emergono con riguardo alle modalità attraverso le quali regolamentare la decisione algoritmica. Una delle opzioni più popolari quando si tratta di tecnologia è quella che poggia sul *risked-based approach* e che, di conseguenza, struttura le regole su più livelli, rendendole più rigorose o più permissive a seconda della posizione occupata sulla scala del rischio.[84] Tuttavia, sebbene l'ipotesi potrebbe apparire *prima facie* ottimale, permane qualche perplessità quando si tenta di calare questo modello ideale nella realtà concreta dei diritti e degli interessi in gioco. In particolare, le scelte che dovranno prendersi con riguardo ai settori ad elevato rischio per le libertà degli individui o alle soglie al di sotto delle quali il rischio viene qualificato come giuridicamente accettabile implicano una graduazione di diritti e di libertà non prevista espressamente dal nostro diritto costituzionale. A questo punto è lecito domandarsi se la forza travolgente dell'algoritmo, che ha già trasformato in maniera radicale la società e l'economia, inaugurerà una stagione di cambiamenti anche per i sistemi costituzionali europei, secondo la logica per cui anche il diritto, al pari di qualsiasi altro organismo vivente, corre il pericolo di soccombere in caso di mancato adattamento ad una realtà circostante in perenne mutamento.[85]

[84] Anche la proposta di regolamento sull'intelligenza artificiale approvata dalla Commissione si fonda su questo approccio.

[85] A. SIMONCINI, S. SUWEIS, *Il cambio di paradigma nell'intelligenza artificiale*, cit.