«Il diritto dell'economia» issn 1123-3036, anno 68, n. 109 (3 2022), pp. 106-138

Articoli e saggi

Nuove sfide della tecnologia e gestione dei rischi nella proposta di regolamento europeo sull'intelligenza artificiale: set di *training,* algoritmi e qualificazione dei dati. Profili critici*

Danila Iacovelli, Marco Fontana

Sommario: 1. Premessa. – 2. Big data e intelligenza artificiale: potenzialità e rischi. – 3. La proposta di regolamento europeo sull'intelligenza artificiale. – 4. I sistemi di intelligenza artificiale ad alto rischio. – 5. Quali dati? Tipologia di dati e algoritmi intelligenti. – 6. La "neutralità dei dati" nella proposta di regolamento e l'eccezione dei dati biometrici. – 7. Dati biometrici tra GDPR e regolamento sull'IA: due esempi concreti. – 8. Conclusioni.

1. *Premessa*

L'articolo esamina alcuni aspetti della proposta di regolamento europeo in materia di intelligenza artificiale ("IA"), con riguardo alla gestione dei rischi per i diritti e le libertà individuali, scaturenti dall'uso delle nuove tecnologie, evidenziando in particolare le problematiche connesse al processo decisionale automatizzato e alla definizione del set di dati di addestramento della macchina, che influisce sulla qualità degli *output* generati dal sistema di IA e ne riflette carenze e parzialità. Le caratteristiche quantitative e qualitative dei dati, tanto quanto le caratteristiche dell'algoritmo, rappresentano una condizione essenziale per il (corretto) funzionamento dell'IA. Se, in base all'impostazione del GDPR, la distinzione più rilevante è quella tra dati "personali" e "non personali", le tecnologie dell'intelligenza artificiale e l'uso dei *big data* mettono spesso in crisi tali separazioni.

* L'articolo è espressione di una riflessione congiunta svolta dagli autori. I paragrafi 1, 2, 3 e 4 sono stati elaborati da Danila Iacovelli; i paragrafi 5, 6, 7 e 8 da Marco Fontana.

La proposta di regolamento è di per sé "neutrale" rispetto ai dati trattati, che continuano ad essere disciplinati dai principi fondamentali del GDPR, mentre definisce in modo dettagliato i sistemi di IA che si basano su dati biometrici, i quali costituiscono i primi sensori dei rischi legati all'uso delle nuove tecnologie. Dal testo del regolamento emergono una serie di limitazioni nell'identificazione delle tecnologie biometriche incluse nel relativo campo di applicazione, che creano, come si evince da alcuni casi concreti considerati, ambiguità e incertezze: queste si riflettono sulla protezione dei diritti individuali e inducono a una riflessione sull'opportunità di adottare scelte più nette in merito alle tecnologie vietate e a quelle ad alto rischio, e di delineare un più chiaro coordinamento con altre fonti normative, a cominciare dal GDPR.

2. *Big data e intelligenza artificiale: potenzialità e rischi*

In un mondo in cui ognuno può condividere la propria esperienza online, attraversando ogni barriera spaziale, per entrare in una rete di interconnessione globale, che diventa fonte di informazione, luogo di incontro (*matchmaker)*, custode delle memorie, modalità di lavoro (*smart working*), lasciamo ogni giorno dietro di noi una serie di tracce, sulle nostre attività, spostamenti, preferenze, sul nostro stesso modo di essere.

I dati non svaniscono. Se moltiplichiamo le piccole tracce per milioni di utenti, ci rendiamo conto dell'enorme mole di dati che può essere raccolta.

L'informazione digitale è radicata in tutti gli aspetti della società e la produzione di dati sembra crescere in maniera inarrestabile[1].

Un recente studio condotto dal fisico Melvin Vopson, presso l'Università di Portsmouth, ipotizza che i dati possano avere una massa fisica misurabile e propone un esperimento per dimostrare se l'informazione possa costituire un quinto stato della materia, oltre a quello solido, liquido, aeriforme, plasmatico[2].

[1] Cfr. M.M. Vopson, *The world's data explained: How much we're producing and where it's all stored*, *World Economic Forum*, 2021, disponibile alla pagina https://www.weforum.org/agenda/2021/05/world-data-produced-stored-global-gb-tb-zb/, in cui si rileva che «Each day on Earth we generate 500 million tweets, 294 billion emails and 4 million gigabytes of Facebook data. Around 150 years from now, the number of digital bits would reach an impossible value, exceeding the number of all atoms on Earth».

[2] M. V. Vopson, *Experimental protocol for testing the mass-energy-information equivalence principle*, in *AIP Adv*, 12, 035311, 2022, 1 ss. Alla base dell'indagine, l'A. pone il principio di equivalenza massa-energia-informazione, esposta in M.M. Vopson, *The mass-energy-information equivalence principle, AIP Adv.*, 095206, 2018, 9, 1 ss. Secondo tale tesi, per esempio, «the data storage device should be heavier when information is stored on it than when it is in fully erased state».

L'idea di un collegamento tra informazione e mondo fisico, benché allo stato non dimostrata, appare senz'altro suggestiva, anche per la connotazione dell'informazione come elemento fondamentale del cosmo.

La raccolta di dati esprime, tuttavia, il suo pieno potenziale se collegata all'uso dell'intelligenza artificiale ("IA") e, a sua volta, ne alimenta la diffusione.

"*Big data*"[3] e "intelligenza artificiale"[4] sono due concetti chiave che contrassegnano l'evoluzione del mondo digitale. I dati di per sé assumono un significato relativo, ma collegati tra loro, analizzati e rielaborati da algoritmi, possono fornire modelli predittivi, utilizzabili per anticipare le risposte degli individui, e quindi anche influenzarne il comportamento per finalità diverse (di marketing, elettorali).

L'intelligenza artificiale si sviluppa e si evolve attingendo alla disponibilità di una enorme quantità di dati, unita al miglioramento della tecnologia e della connettività, nonché alla capacità di imparare dai dati stessi per generare nuovi contenuti[5], secondo un processo di apprendimento automatico[6].

Le applicazioni sono numerose e ricoprono una molteplicità di settori: dal *natural language processing* (NLP); al riconoscimento vocale; al *data mining*; all'elaborazione grafica e delle immagini; alla guida autonoma; alla robotica; al sistema di riconoscimento biometrico; agli sviluppi della biotecnologia, sempre più connes-

[3] Il termine "Big data", pur identificato in vari modi, si collega a tre elementi rilevanti, definiti 3V di crescita: volume, varietà, velocità. Cfr. D. E. O'Leary, *Artificial Intelligence and Big Data,* in *IEEE Intelligent Systems*, 2013, 28, 2, 96 ss., in cui si rileva che la connotazione attraverso una o tutte le tre "V" è forse la più conosciuta, specificando che «*Volume* refers to larger amounts of data being generated from a range of sources. For example, big data can include data gathered from the Internet of Things (IoT) […] *Variety* refers to using multiple kinds of data to analyze a situation or event. […] *Velocity* of data also is increasing rapidly over time for both structured and unstructured data, and there's a need for more frequent decision making about that data».

[4] Cfr. Comunicazione della Commissione europea del 4.3.2019, COM(2018) 795, avente ad oggetto Piano coordinato per lo sviluppo e l'utilizzo dell'intelligenza artificiale, in cui si evidenzia che «Per "intelligenza artificiale" (IA) si intendono quei sistemi che mostrano un comportamento intelligente analizzando il proprio ambiente e compiendo azioni, con un certo grado di autonomia, per raggiungere obiettivi specifici. Usiamo l'IA quotidianamente, ad esempio per bloccare lo spam nella posta elettronica o per parlare con gli assistenti digitali. L'aumento della potenza di calcolo e della disponibilità dei dati e il progresso negli algoritmi hanno reso l'IA una delle tecnologie più importanti del 21° secolo».

[5] Cfr. Comunicazione della Commissione europea del 26.6.2018, COM(2018) 237, avente per oggetto «L'intelligenza artificiale per l'Europa», 10 ss., in cui si evidenzia che «Sono necessari ingenti volumi di dati per sviluppare l'IA. L'apprendimento automatico, un tipo di IA, opera mediante l'individuazione di modelli a partire dai dati disponibili e la successiva applicazione di questa conoscenza ai dati nuovi. Quanto più è grande il set di dati, tanto più accurata sarà l'individuazione delle relazioni anche impercettibili tra i dati».

[6] Cfr. F. Mattassoglio, *Algoritmi e regolazione. Circa i limiti del principio di neutralità tecnologica*, in *Riv. Regolazione dei mercati,* 2018, 2, 231 ss., in cui si specifica che «Questo tipo di elaborazione comprende di solito almeno due operazioni parallele, ossia due tipi di algoritmi diversi: il *classifier* e il *learner*. Il primo si occupa di elaborare i dati e produrre, su quella base, un determinato risultato. Il secondo, invece, testa questi risultati al fine di attribuire loro una matrice del peso che sarà poi in seguito utilizzata dal primo algoritmo per classificare nuovi dati e situazioni che i programmatori non avevano considerato all'inizio».

si all'uso dei *big data*, che offrono enormi opportunità e prospettive per la salute umana, oltre ad agevolare risposte più efficaci alle emergenze sanitarie pubbliche[7].

Il fulcro del sistema è l'algoritmo, che «oggi, per lo più, identifica un software costituito, a sua volta, dai c.d. codici sorgenti o linguaggio sorgente, che rappresentano le istruzioni, fornite dagli operatori umani per eseguire una certa funzione o risolvere un determinato problema. Dette istruzioni possono essere assimilate a scatole (box) nelle quali vengono inseriti i dati necessari a ottenere i risultati voluti»[8].

Oltre ad algoritmi semplici, che utilizzano determinate categorie di dati e operano secondo il modello impostato dal relativo ideatore – con esiti prevedibili (c.d. scatola bianca), o facilmente interpretabili (c.d. scatola grigia) –, vi sono una serie di algoritmi complessi, che evolvono attraverso un processo di *machine learning*, secondo logiche che possono risultare incomprensibili, nonché algoritmi "senzienti", in grado di superare il cosiddetto test di Turing[9], ovvero di equiparare il funzionamento della mente umana[10].

Il problema più critico si pone quando l'apprendimento delle macchine rende difficile, o impossibile, capire cosa stiano imparando e quali decisioni stiano adottando (*Black Box Dilemma*)[11]. Le scatole nere si suddividono in due categorie: quelle forti (*strong black boxes*), in cui il comportamento dell'IA è totalmente opaco per l'essere umano, che non è in grado di comprendere il processo decisionale, né la classificazione delle variabili e i risultati prodotti; e quelle deboli (*weak black boxes*), che consentono una limitata capacità di speculazione su come l'IA sia pervenuta a determinate decisioni[12].

---

[7] Cfr. J.T. O'Brien, C. Nelson, *Assessing the Risks Posed by the Convergence of Artificial Intelligence and Biotechnology*, in *Health Security*, 2020, 18, 3, 219 ss. ; A.L. Oliveira, *Biotechnology, Big Data and Artificial Intelligence*, in *Biotechnology Journal*, 2019, 14, 8.

[8] F. Mattassoglio, *Algoritmi e regolazione. Circa i limiti del principio di neutralità tecnologica*, cit., 230.

[9] Il Test di Touring consente di verificare se una macchina si in grado di avere un comportamento intelligente, ed è stato proposto da A.M. Turing, nell'articolo *Computing Machinery and Intelligence*, in *Mind*, 1950, 59, 433 ss.

[10] Cfr. L. Nalbandian, *Using Machine-Learning to Triage Canada's Temporary Resident Visa Application,* Working papers, 2021/9, produced jointly by the Ryerson Centre for Immigration and Settlement (RCIS) and the CERC, in https://km4s.ca/wp-content/uploads/2021_9_Nalbandian_Lucia_Using_Machine_Learning _to_Triage_Canadas_ Temporary_Resident_Visa_Applications.pdf, 2021, 3, in cui si precisa che «Deep learning is a branch of machine learning that allows computers to perform human tasks, like understanding a language or identifying images, which uses algorithms inspired by the structure and function of the human brain (Brownlee, 2019 [*What is Deep Learning? Machine Learning Mastery*, https://machinelearningmastery.com/what-is-deep-learning/] Deep learning employs "artificial neural networks" to undertake both supervised and unsupervised learning methods. In the human brain, neural networks are made up of neurons that process information received from our senses to deliver an output or response (Panchal, 2018) [*Artificial Neural Networks – Mapping the Human Brain. Medium,* https://medium.com/predict/artificial-neural-networks-mapping-the-human-brain2e0bd4a9316].

[11] In arg., ivi, 4.

[12] *Ibidem*.

Uno dei settori che trae indubbi vantaggi dall'uso dell'IA è quello della finanza tecnologica, in quanto i nuovi algoritmi consentono di segmentare la clientela, sulla base dell'analisi dei relativi comportamenti e preferenze e quindi orientare le strategie di *marketing*, adeguandole alle specifiche esigenze dei destinatari dei servizi finanziari. Il procedimento di profilazione consente alle imprese di definire precisamente i prodotti da immettere sul mercato, ottimizzando i modelli di *business* e i costi[13], ma al tempo stesso comporta una serie di rischi per la tutela della *privacy* e delle libertà individuali, tanto da richiedere – come si dirà oltre – una regolamentazione specifica.

L'intreccio tra IA e *big data* può produrre impatti che si estendono ben al di là di un particolare settore, per influenzare le sorti di intere società. È il caso, ormai emblematico, di Cambridge Analytica[14], la società che, utilizzando i dati degli utenti di Facebook, raccolti attraverso l'applicazione "mydigitalife"[15], è stata in grado di manipolare l'opinione pubblica in vista delle elezioni statunitensi, al punto che ci si domanda se elezioni libere ed eque siano ancora possibili[16].

---

[13] Cfr. F. Mattassoglio, *Algoritmi e regolazione. Circa i limiti del principio di neutralità tecnologica*, cit., 235, in cui si rileva, con specifico riferimento al settore finanziario, che «Le più ampie informazioni sul cliente, inoltre, possono consentire alle imprese di definire con maggior precisione il giusto prezzo dei prodotti finanziari, per evitare pericolosi fenomeni di *adverse selection* e *free riding*. Nell'ambito della valutazione del rischio di credito, ad esempio, recenti studi hanno messo in luce come il ricorso a tecniche di BDA possa migliorare l'accuratezza dei giudizi, utilizzando nuovi modelli di reti connesse (*correlation network*) che si riferiscono al soggetto richiedente aumentando le informazioni a disposizione del finanziatore». Sul tema, tra gli altri, L. Ammannati, G.L. Greco, *Piattaforme digitali, algoritmi e "big data": il caso del "credit scoring"*, in *Rivista Trimestrale di Diritto dell'Economia*, 2021, 2, 1, 290 ss.; P. Giudici, G. Polinesi, *Scoring models for robo-advisory platforms: a network approach,* in *Journal of Network Theory in Finance*, 2020, 5, 2, 1 ss; M. Zappatore, *Big Data, Profilazione e Mercati Finanziari: utilizzo e tutela*, in *Giuricivile,* 4, 2019 (https://giuricivile.it/big-data-profilazione-e-mercati-finanziari-utilizzo-e-tutela/); R. Motroni, *La classificazione della clientele nella normativa dei mercati e degli strumenti finanziari,* in Giureta, *Rivista di Diritto dell'Economia, dei Trasporti e dell'Ambiente*, 2015, 13, 425 ss.

[14] Cambridge Analytica era una filiale di SCL Group (Group Strategic Communication Laboratories), fondata nel 2013, che è stata in grado di influenzare la campagna presidenziale negli Stati Uniti d'America, e ha assunto un ruolo significativo nella Brexit. La società britannica ha raccolto illecitamente i dati di oltre 87 milioni di utenti Facebook, senza il loro consenso, e li ha utilizzati per scopi di propaganda politica. Il sistema utilizzato da Cambridge Analytica si basa sul cosiddetto "*microtargeting* psicografico", che consiste nel valutare la personalità degli utenti online attraverso i dati personali lasciati su internet e quindi influenzarne le opinioni e le scelte attraverso l'invio di messaggi pubblicitari altamente personalizzati. La società statunitense era riuscita a creare un algoritmo, basato su un modello elaborato dal ricercatore e psicologo Michal Kosinski, che consentiva di agire non solo in base alle preferenze degli utenti, come nel marketing tradizionale, ma anche sulla loro parte emotiva, anticipandone le reazioni. Lo scandalo di Cambridge Analytica è emerso a seguito delle rivelazioni del *The Guardian* e del *New York Times*, nel 2018. Si veda l'articolo di C. Cadwalladr, E. Graham-Harrison, *Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach*, in https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election.

[15] La Facebook app creata da Alexander Kogan, che ha consentito di collezionare i dati degli utenti che l'hanno scaricata e quelli dei loro amici, pervenendo ad una raccolta di oltre 50 milioni di informazioni.

[16] Così, la giornalista Cadwalladr (v. *supra* nt. 14), in un noto talk sul tema, tenutosi nell'aprile 2019: *Facebook's role in Brexit – and the threat to democracy,* in https://www.ted.com/talks/carole_cadwalladr_facebook_s_role_in_brexit_and_the_threat_to_democracy.

Una raccolta massiva di dati associati all'uso dell'intelligenza artificiale può incidere sensibilmente sui diritti e le libertà delle persone, non solo nell'ambito di processi consapevolmente mirati a influenzare il comportamento degli individui, dei consumatori, e perfino le decisioni di una collettività, ma anche per effetto di rischi intrinseci al funzionamento dell'IA.

Simili rischi possono verificarsi, ad esempio, quando si affidano all'IA processi decisionali tipicamente umani, con esiti difficilmente prevedibili *a priori*, anche in dipendenza delle funzioni di addestramento della macchina, che possono condurre a risultati fortemente pregiudizievoli per interi gruppi di persone.

La letteratura[17] ha individuato diverse fasi della programmazione e del successivo funzionamento della macchina suscettibili di produrre effetti potenzialmente discriminatori, quali l'individuazione delle "*target variables*" e delle "*class labels*"; il *training dei dati*; la selezione delle funzionalità del modello impostato dal programmatore ("*feature selection*")[18]; i "*proxies*"[19], oltre al caso della discriminazione volontaria, dovuta alla selezione dei dati che si riverbera su determinate classi di persone.

Un caso sintomatico del rischio di distorsioni discriminatorie è stato quello dell'implementazione di sistemi decisionali automatizzati nel contesto canadese del diritto all'immigrazione dei rifugiati. Si tratta di un contesto rilevante, sia per la specifica vulnerabilità, in quanto le comunità dei non cittadini ricevono spesso una tutela dei diritti meno solida e dispongono di risorse più limitate a protezione dei propri diritti[20], sia per la grande portata del relativo impatto[21].

---

[17] Cfr. S. Barocas, A.D. Selbs, *Big data disparate impac*t, in *California Law Review*, 2016, 671 ss. Sul tema, C. Nardocci, *Intelligenza artificiale e discriminazioni*, Convegno annuale dell'associazione "Gruppo di Pisa" Il diritto costituzionale e le sfide dell'innovazione tecnologica 18 e 19 giugno 2021, in https://www.gruppodipisa.it/images/convegni/2021_Convegno_Genova/Costanza_Nardocci_-_Intelligenza_artificiale_e_discriminazioni.pdf, 18.

[18] Cfr. Barocas, A.D. Selbs, *Big data disparate impac*t, cit., 688, in cui si osserva che «Through a process called "feature selection," organizations – and the data miners that work for them – make choices about what attributes they observe and subsequently fold into their analyses. These decisions can also have serious implications for the treatment of protected classes if those factors that better account for pertinent statistical variation among members of a protected class are not well represented in the set of selected features».

[19] Ivi, p. 691, in cui si rileva che «Cases of decision making that do not artificially introduce discriminatory effects into the data mining process may nevertheless result in systematically less favorable determinations for members of protected classes. This is possible when the criteria that are genuinely relevant in making rational and well informed decisions also happen to serve as reliable proxies for class membership. In other words, the very same criteria that correctly sort individuals according to their predicted likelihood of excelling at a job – as formalized in some fashion – may also sort individuals according to class membership».

[20] P. Molnar, L. Gill, *Bots at the gate, A human rights analysis of automated decision-making in Canada's immigration and refugee system,* University of Toronto's International Human Rights Program (IHRP), Faculty of Law, University of Toronto, and the Citizen Lab (Munk School of Global Affairs and Public Policy, University of Toronto), in https://citizenlab.ca/wp-content/uploads/2018/09/IHRP-Automated-Systems-Report-Web-V2.pdf, 2018, 4.

[21] Ivi*,* in cui si rileva che «The sheer scale of potential impact is another reason for the focus on the immigration and refugee context. In 2017, Immigration, Refugees and Citizenship Canada (IRCC) and the

L'esperienza canadese ha rappresentato «un laboratorio ad alto rischio» nel processo decisionale automatizzato[22]. Per i richiedenti asilo, le conseguenze di una domanda respinta possono sfociare in persecuzioni, sulla base della razza, delle opinioni politiche, del credo religioso, o dell'appartenenza a determinati gruppi etnici.

In tale quadro, il Canada sin dal 2014 ha sviluppato un sistema di analisi predittiva per automatizzare le attività condotte dall'amministrazione in questo settore, al fine di identificare il merito delle domande, individuare segnali di allarme per frode e valutare il complesso dei fattori che incidono sull'accoglimento delle domande[23].

Nel 2018, il Canada ha anche adottato un Libro bianco, avente ad oggetto «*Responsible Artificial Intelligence in the Government of Canada*», che prevede lo sviluppo di strumento adeguati, affinché le istituzioni possano valutare il livello di automazione appropriato rispetto ai relativi programmi e i rischi potenziali per i diritti e gli interesse delle persone, ponendo l'accento su alcuni principi, quali una *governance* democratica, la responsabilità e trasparenza, e sulla necessità che i sistemi implementati siano addestrati a riflettere gli obblighi in materia di diritti umani, ma si tratta di principi non vincolanti. Contestualmente, il *Treasury Board of Canada Secretariat* (TBCS) ha avviato una consultazione per una proposta di direttiva, che richiederebbe una valutazione di impatto algoritmica al momento dell'adozione di nuove tecnologie, e una valutazione retroattiva per quelle già in uso[24]. Tuttavia, anche tali principi non sono vincolanti.

La complessità della migrazione e dello status di rifugiato non è facilmente riconducibile ad un algoritmo.

Canada Border Services Agency (CBSA) processed over 50,000 refugee claims. Canada is projecting the admission of 310,000 new permanent residents in 2018 and up to 340,000 new permanent residents annually by 2020. In 2016, there were over 266,000 students holding an international study permit in Canada; the government issued over 10,000 Temporary Resident Permits (TRPs) and extensions; it approved over 1.3 million applications for individuals wanting to visit Canada and over 2.5 million electronic travel authorization (ETA) applications; and it issued over 78,500 work permits under the Temporary Foreign Worker (TFW) Program and over 208,500 work permits under the International Mobility Program (IMP). Cumulatively, these programs involve millions of determinations and decisions every year. They change the course of millions of lives». Sul punto, L. Nalbandian, cit., 6, in cui si ricorda che «As per section 22 of the Immigration and Refugee Protection Act (IRPA), Temporary resident 22 (1) A foreign national becomes a temporary resident if an officer is satisfied that the foreign national has applied for that status, has met the obligations set out in paragraph 20(1) (b), is not inadmissible and is not the subject of a declaration made under subsection 22. Obligation on entry 20 (1) Every foreign national, other than a foreign national referred to in section 19, who seeks to enter or remain in Canada must establish, […] (b) to become a temporary resident, that they hold the visa or other document required under the regulations and will leave Canada by the end of the period authorized for their stay».

[22] P. Molnar, L. Gill, cit., 4.

[23] Ivi, 14.

[24] Ivi, 16 ss., in cui si precisa che «The inclusion of AIAs into this framework is motivated by calls in 2018 from Nesta and the AI Now Institute, two major civil society agencies based in the United Kingdom and the United States respectively».

L'uso dell'intelligenza artificiale investe ogni aspetto del vivere sociale, sotto il profilo economico, politico, sanitario, ambientale[25], e non si limita ai modelli aziendali o ai sistemi di gestione dell'amministrazione pubblica, ma ingenera una serie di trasformazioni nelle abituali interazioni con il mondo esterno, a partire dal contesto abitativo[26] e urbano[27].

Tali cambiamenti implicano la necessità di valutare adeguatamente i benefici connessi all'impiego di nuove tecnologie e i rischi per le persone fisiche e la società, in una logica di compatibilità che garantisca uno sviluppo affidabile e sicuro dell'intelligenza artificiale e la salvaguardia dei principi etici, come richiesto dal Consiglio e dal Parlamento europeo[28].

L'esigenza di armonizzare la protezione dei diritti e delle libertà fondamentali delle persone fisiche e assicurare la libera circolazione dei dati tra gli stati membri è stata avvertita in primo luogo rispetto all'attività di trattamento dei dati e posta alla base del Regolamento (UE) 2016/679 ("GDPR")[29]. L'articolo 8, paragrafo 1,

[25] Come evidenziato nella comunicazione della Commissione europea COM(2018) 237, cit., «Oltre a renderci più facile la vita, l'IA ci aiuta a risolvere alcune tra le sfide più ardue al mondo: dal trattamento delle malattie croniche o dalla riduzione dei tassi di incidenti stradali mortali, alla lotta contro il cambiamento climatico o alla prevenzione delle minacce alla sicurezza informatica». Sul tema, F. Pizzetti, *La protezione dei dati personali dalla direttiva al nuovo regolamento: una sfida per le Autorità di controllo e una difesa per la libertà dei moderni*, in *MediaLaws*, 1, 2018, 123, in cui si osserva che «Viviamo in un mondo nel quale sempre più la dimensione della realtà digitale e quella della realtà reale si intrecciano fra loro mentre la prospettiva di una civiltà dominata dai Big Data, dalla Intelligenza Artificiale, dall'Internet delle cose, è già nel nostro presente e ancor più nel nostro prossimo futuro».

[26] Si pensi, in generale, all'*Internet of things* o, in particolare, per esempio, ai termostati intelligenti, che analizzano le abitudini di chi vive in un determinato ambiente e regolano la temperatura in modo efficiente.

[27] Per alcuni efficaci esempi, si veda F. Meneghetti, C. Rossi Chauvenet, G. Fioroni, *Rapporto 3/2022-SMART cities e intelligenza artificiale*, in *BioLaw Journal – Rivista di BioDiritto*, 2022, 1, 253 ss., in cui si fa riferimento alla creazione di una piattaforma di Urban intelligence nel comune di Venezia, nell'ambito di un partenariato per l'innovazione. In esito alla raccolta di tera di informazioni sulla città, è possibile effettuare, «tra le altre cose, una analisi dei flussi (ad esempio, quante persone sono entrate nel Comune di Venezia a partire da un determinato orario). […] Un altro esempio attiene alla pedonabilità dell'isola, grazie a circa una cinquanta sensori distribuiti all'interno della città, che fanno analisi del traffico pedonale tramite machine learning. Ed ancora, circa 80 telecamere sono disposte sui canali della città di Venezia per osservare il traffico acqueo, analizzato tramite intelligenza artificiale che consente di classificare le imbarcazioni per tipologia». Analoghe piattaforme possono essere organizzate in ambiti diversi come il turismo, la sicurezza, la mobilità, la transizione ambientale».

[28] Cfr. Consiglio europeo, Riunione straordinaria del Consiglio europeo (1° e 2 ottobre 2020) - Conclusioni, EUCO 13/20, 2020, p. 7; Parlamento europeo, Risoluzione del 20 ottobre 2020 recante raccomandazioni alla Commissione concernenti il quadro relativo agli aspetti etici dell'intelligenza artificiale, della robotica e delle tecnologie correlate (2020/2012(INL)).

[29] Reg. (UE) 2016/679 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 Dir. 95/46/CE». Si veda, special., art. 3 del regolamento. Sul tema, in un'ampia letteratura, G. Finocchiaro (a cura di), *La protezione dei dati personali in Italia. Regolamento UE 2016/679 e d.lgs. 10 agosto 2018, n. 101*, Bologna, 2019; R. Panetta (a cura di), *Circolazione e protezione dei dati personali, tra libertà e regole del mercato. Commentario al Regolamento UE n. 2016/679 (GDPR) e al novellato d.lgs. n. 196/2003 (Codice privacy)*, Milano, 2019; E. Tosi (a cura di), *Privacy digitale. Riservatezza e protezione dei dati personali tra GDPR e nuovo Codice Privacy*, Milano, 2019; C. Colapietro, *I principi ispiratori del Regolamento UE 2016 sulla protezione dei dati personali e la loro incidenza sul contesto normativo nazionale*, in *federalismi.it,* 21.11.2018; P.G. Alpa, *L'identità digitale e la tutela della per-*

della Carta dei diritti fondamentali dell'Unione europea («Carta») e l'articolo 16, paragrafo 1, del trattato sul funzionamento dell'Unione europea («TFUE») stabiliscono che ogni persona ha diritto alla protezione dei dati di carattere personale che la riguardano, come pur ricordato nel primo considerando del GDPR.

I dati sono la risorsa fondamentale per lo sviluppo dell'intelligenza artificiale, quale «mezzo attraverso il quale l'algoritmo apprende e interagisce con il suo ambiente»[30], e il regolamento sulla protezione dei dati, come rilevato dalla Commissione europea, rappresenta una «tappa fondamentale per costruire la fiducia»[31] e la base per sviluppare un approccio sostenibile alle nuove tecnologie, in linea con la strategia per un mercato unico digitale[32].

La politica europea sulla protezione dei dati personali svolge un ruolo centrale per la responsabilizzazione nell'utilizzo delle nuove tecnologie e per la creazione di un ambiente di sicurezza informatica, attraverso la definizione di un contesto normativo affidabile e la predisposizione di una serie di misure di salvaguardia che assicurino la protezione dei diritti e delle libertà fondamentali dei cittadini e delle imprese nell'ambito dell'Unione europea.

L'evoluzione dell'intelligenza artificiale e l'innesto con i *big data* hanno tuttavia imposto una riflessione ulteriore sulla sostenibilità del quadro normativo vigente e sulle questioni relative alla responsabilità civile[33].

La Commissione europea[34] ha quindi avviato un percorso di valutazione dell'adeguatezza dell'attuale sistema di norme rispetto alle nuove sfide e di verifica delle possibili lacune, evidenziando l'importanza di implementare la sensibilizzazione sugli algoritmi[35].

---

sona. Spunti di riflessione, in *Contr. Impr.*, 2017, 3, 723 ss.; L. Califano, *Il Regolamento UE 2016/679 e la costruzione di un modello uniforme di diritto europeo alla riservatezza e alla protezione dei dati personali*, in L. Califano, C. Colapietro (a cura di), *Innovazione tecnologica e valore della persona. Il diritto alla protezione dei dati personali nel Regolamento UE 2016/ 679*, Napoli, 2017, 3 ss.; Id., *Privacy: affermazione e pratica di un diritto fondamentale*, Napoli, 2016; F. Pizzetti, *Privacy e il diritto europeo alla protezione dei dati personali. Il Regolamento europeo 2016/679*, Torino, 2016.

[30] Comunicazione della Commissione europea COM(2018) 237, cit., 11, in cui peraltro si evidenzia che «L'UE si è impegnata notevolmente negli ultimi 15 anni per rendere disponibili le informazioni del settore pubblico e i risultati delle ricerche a finanziamento pubblico per il riutilizzo, come i dati generati dai programmi spaziali dell'UE» e che le decisioni politiche «dovrebbero anche incoraggiare la più ampia disponibilità di dati detenuti a titolo privato, assicurando al contempo il pieno rispetto della legislazione sulla protezione dei dati di carattere personale».

[31] Ivi, 2.

[32] Comunicazione della Commissione europea del 6.5.2015, COM(2015) 192, avente per oggetto «Strategia per il mercato unico digitale in Europa»; Comunicazione della Commissione europea del 10.5.2017, COM(2017) 228, avente per oggetto «Un mercato unico digitale connesso per tutti».

[33] Cfr. Comunicazione della Commissione europea COM(2018) 237, cit., 17.

[34] *Ibidem*.

[35] Ivi, 18, in cui si evidenzia che la Commissione provvederà, tra l'altro, a «sostenere la ricerca sullo sviluppo dell'IA spiegabile e implementare un progetto pilota proposto dal Parlamento europeo riguardante la sensibilizzazione sugli algoritmi, per raccogliere una solida base di prove a sostegno dei progetti di risposte poli-

L'esigenza di una regolazione è stata avvertita, ancor prima, in alcuni settori specifici, sia al livello nazionale che sovranazionale. In particolare, il legislatore italiano è intervenuto con alcune disposizioni per definire la *blockchain* e gli *smart contracts,* nell'ambito di una normativa di emergenza[36], con un metodo pur criticato in dottrina in quanto la norma «dovrebbe essere tecnologicamente neutra e dunque non riferirsi ad una particolare tecnologia, affermata in un particolare momento storico»[37].

In base al principio di neutralità tecnologica, o principio dell'"open Internet", sancito dal regolamento (UE) 2015/2120[38], l'accesso a Internet deve essere trattato in modo non discriminatorio, indipendentemente dal contenuto, dall'applicazione, dal servizio, dal terminale, nonché dal mittente e destinatario[39].

Tuttavia, la rapida evoluzione di alcuni settori, che si avvalgono di tecnologie altamente complesse, come quello finanziario, già sopra ricordato, ha posto la necessità, sul piano europeo e quindi su quello nazionale, di adottare una regolamentazione specifica[40].

---

tiche alle problematiche indotte dal processo decisionale automatizzato, tra cui i condizionamenti e le discriminazioni». Sul progetto pilota: https://ec.europa.eu/digital-single-market/en/algorithmic-awareness-building.

[36] Decreto-legge 14 dicembre 2018, n. 135, convertito in legge 11 febbraio 2019, n. 12, art. 8 *ter*.

[37] G. Finocchiaro, *Intelligenza artificiale e protezione dei dati personali*, in *Giur. it.,* 2019, 1670 ss, in cui si osserva che «L'allarme sociale, assolutamente condivisibile in molti casi, produce quasi automaticamente la richiesta di norme *ad hoc*. Mentre è comprensibile la richiesta di una sicura dichiarazione di illiceità del fenomeno con le eventuali conseguenze sotto il profilo sanzionatorio, l'esigenza di prevedere una norma specifica per ogni nuova fattispecie appare l'esito di una carenza metodologica». In senso contrario, per cui la regolazione delle nuove tecnologie (e in particolare del regime di responsabilità civile che le riguarda) dovrebbe basarsi su un approccio "neutrale" ma con soluzioni specifiche per ogni "classe" di tecnologia, cfr. il già citato F. Mattassoglio, cit., 229 ss. e A. Bertolini, *Artificial intelligence and civil liability*, studio per Policy Department for Citizens' Rights and Constitutional Affairs, European Parliament, Directorate-General for Internal Policies, Bruxelles, 2020, 15 (disponibile alla pagina https://www.europarl.europa.eu/RegData/etudes/STUD/2020/621926/IPOL_STU(2020)621926_EN.pdf). V. anche, con specifico riguardo al settore dei veicoli a guida autonoma, A. Bertolini, M. Riccaboni, *Grounding the case for a European approach to the regulation of automated driving: the technology-selection effect of liability rules*, in *European Journal of Law and Economics*, 2021, 51, 248.

[38] Regolamento (UE) 2015/2120, «che stabilisce misure riguardanti l'accesso a un'Internet aperta e che modifica la direttiva 2002/22/CE relativa al servizio universale e ai diritti degli utenti in materia di reti e di servizi di comunicazione elettronica e il regolamento (UE) n. 531/2012 relativo al roaming sulle reti pubbliche di comunicazioni mobili all'interno».

[39] Su punto, Autorità per le garanzie nelle telecomunicazioni ("AGCOM"), *Internet aperta /Neutralità della rete*, in *www.agcom.it*, che evidenzia come in tale contesto «Le autorità nazionali di regolamentazione sono tenute a promuovere la disponibilità dell'accesso a Internet a livelli qualitativi che siano al passo con il progresso tecnologico e in maniera non discriminatoria. A tal fine, possono imporre agli operatori di comunicazione elettronica e ai fornitori di servizi di accesso a Internet requisiti tecnici di qualità del servizio e altre misure adeguate e necessarie».

[40] Cfr. F. Mattassoglio, *Algoritmi e regolazione. Circa i limiti del principio di neutralità tecnologica*, cit., 237 ss., dove si individuano una serie di settori in cui «si può cogliere la tendenza alla regolazione e al controllo del procedimento algoritmico», quali «la negoziazione algoritmica ad alta frequenza (High Frequency Trading o HFT) 79 – intesa come sottoinsieme della più ampia categoria di trading che poggia su ordini gestiti in modo automatico – in cui le sequenze sono elaborate ad altissima velocità (es. 7 mila operazioni in un istante)», in relazione alla quale è stata emanata la direttiva 2014/65/UE, relativa ai mercati degli strumenti finanziari e che modifica la direttiva 2002/92/CE e la direttiva 2011/61/UE, che in particolare all'art. 17 dispone che «Le

In tale scenario, considerata l'esigenza di promuovere il progresso e lo sviluppo dell'IA e salvaguardare, al tempo stesso, i diritti e le libertà delle persone, la Commissione europea ha elaborato una proposta di regolamento «che stabilisce regole armonizzate sull'intelligenza artificiale e modifica alcuni atti legislativi dell'Unione»[41], con l'obiettivo di adottare «una normativa per un approccio europeo coordinato alle implicazioni umane ed etiche dell'intelligenza artificiale».

3. *La proposta di regolamento europeo sull'intelligenza artificiale*

La proposta di regolamento è stata preceduta dall'adozione di un Libro Bianco sull'intelligenza artificiale[42], nel quale si delinea una strategia volta a conciliare lo sviluppo sicuro e affidabile dell'IA con il «pieno rispetto dei valori e dei diritti dei cittadini dell'UE», in modo da creare un «ecosistema di fiducia»[43], secondo un approccio antropocentrico: «l'intelligenza artificiale non è fine a se

imprese di investimento che effettuano negoziazione algoritmica pongono in essere controlli dei sistemi e del rischio efficaci e idonei per l'attività esercitata volti a garantire che i propri sistemi di negoziazione siano resilienti e dispongano di sufficiente capacità, siano soggetti a soglie e limiti di negoziazione appropriati e impediscano l'invio di ordini erronei o comunque un funzionamento dei sistemi tale da creare un mercato disordinato o contribuirvi». Un altro ambito in cui si è diffuso il ricorso ad algoritmi complessi è, ad esempio, quello della c.d. consulenza finanziaria automatizzata (o *roboadvice),* in cui la componente algoritmica, finalizzata alla profilazione, è stata considerata quale mera modalità alternativa di offerta del tradizionale servizio di consulenza, in apparente coerenza col principio di neutralità tecnologica. Ad essa, infatti, si ritiene debbano essere per lo più applicate le regole generali oggi contenute negli artt. 21 e 24-*bis* TUF, relative agli obblighi di informazione preventiva del cliente e di condotta in capo all'intermediario. Tuttavia, l'ESMA ha rilevato i rischi legati all'automazione del servizio di consulenza, come emerge dalle recenti linee guida sui requisiti di adeguatezza per la protezione degli investitori (ESMA, *Final Report*, *Guidelines on certain aspects of the MiFID II suitability requirements,* in https://www.esma.europa.eu/sites/default/files/library/esma35-43-869-_fr_on_guidelines_on_suitability.pdf). Un ulteriore settore in cui si verifica la tendenza alla regolazione dei procedimenti algoritmici è quello del merito creditizio, in cui la Banca Centrale Europea ha introdotto una fase di valutazione sulla *governance* e sul processo decisionale creditizio (BCE, *Guide to assessments of Fintech institution licence applications*, marzo 2018, in *www.bankingsupervision.europa.eu*).

[41] Comunicazione della Commissione europea del 21 aprile 2021, COM(2021) 206 final. Sul tema, tra gli altri, C. Casonato, B. Marchetti, *Prime osservazioni sulla proposta di regolamento dell'Unione europea in materia di intelligenza artificiale,* in *BioLaw Journal – Rivista di BioDiritto*, 2021, 3, 415 ss.; M. Peruzzi, *Il diritto antidiscriminatorio al test di intelligenza artificiale,* in *Labour & Law Issues*, 2021, 1, 48 ss.; G. Marchianò, *Proposta di regolamento della Commissione europea del 21 aprile 2021 sull'intelligenza artificiale con particolare riferimento alle IA ad alto rischio*, in *ambientediritto.it,* 2021, 2, 616 ss.(disponibile su https://www.ambientediritto.it/dottrina/proposta-di-regolamento-della-commissione-europea-del-21-04-2021-sullintelligenza-artificiale-con-particolare-riferimento-alle-ia-ad-alto-rischio/); G. Contissa, F. Gallli, F. Godano, G. Sartor, *Il regolamento europeo sull'intelligenza artificiale. Analisi informatico-giuridica,* in *i-lex.it*, 2022, 1.

[42] Comunicazione della Commissione europea del 19.2.2020, COM(2020) 65 final, avente per oggetto «Libro Bianco sull'intelligenza artificiale Un approccio europeo all'eccellenza e alla fiducia».

[43] Ivi, 3.

stessa, ma è uno strumento a servizio delle persone che ha come fine ultimo quello di migliorare il benessere degli esseri umani»[44].

Pertanto, le applicazioni dell'intelligenza artificiale «dovrebbero non solo rispettare la legge, ma anche osservare i principi etici e garantire che le loro attuazioni pratiche non comportino danni indesiderati»[45], puntando a rafforzare «le capacità dei cittadini, non a sostituirsi a loro, e consentire l'accesso anche alle persone con disabilità»[46].

In tale contesto socio-politico, la proposta di regolamento mira ad istituire un quadro giuridico uniforme tra i paesi membri dell'UE, «solido e flessibile», che garantisca uno sviluppo equilibrato dell'IA, con un approccio incentrato sui rischi, secondo un criterio di proporzionalità, in modo da non creare ostacoli e restrizioni al mercato[47].

L'introduzione di una normativa flessibile è un aspetto nevralgico del regolamento, in una proiezione futura, al fine di tutelare diritti che, per natura, sono esposti a una pluralità di rischi, scaturenti dalla continua e rapida evoluzione della società digitale e della tecnologia, con uno schema che ricalca, sotto questo profilo, l'impianto del GDPR, improntato a un sistema di principi e meccanismi dinamici, destinati a mettersi in relazione a una realtà mutevole[48].

Tali meccanismi ricomprendono, ad esempio, la valutazione di impatto, obbligatoria per una serie di trattamenti che comportino rischi per i diritti e le libertà delle persone, in particolare quando siano utilizzate nuove tecnologie; gli obblighi di consultazione dell'Autorità garante, qualora sussistano rischi residui elevati, secondo uno schema che fa leva sul *principio* di responsabilità del titola-

[44] Comunicazione della Commissione europea del 8.4.2019, COM(2019) 168 final, avente per oggetto «Creare fiducia nell'intelligenza artificiale antropocentrica», 2. Si veda anche, comunicazione della Commissione europea del 7.10.2018, COM(2018) 795 final, contenente un «Piano coordinato sull'intelligenza artificiale», al fine di sviluppare sinergie e promuovere investimenti congiunti.

[45] Comunicazione della Commissione europea COM(2019) 168 final, cit., 2.

[46] Ivi, 2. La Commissione europea, in particolare, ha rilevato che «Vi è quindi la necessità di elaborare orientamenti etici basati sul quadro normativo esistente e che dovrebbero essere applicati da sviluppatori, fornitori e utenti dell'IA nel mercato interno, stabilendo condizioni di parità sul piano etico in tutti gli Stati membri. Per questo motivo la Commissione ha istituito un gruppo di esperti ad alto livello sull'IA in rappresentanza di una vasta gamma di portatori di interessi, incaricandolo della stesura di orientamenti etici sull'IA e della redazione di una serie di raccomandazioni per una più ampia politica in materia di IA. Allo stesso tempo è stata istituita l'Alleanza europea per l'IA, piattaforma multilaterale aperta con oltre 2 700 membri, per fornire un contributo più vasto ai lavori del gruppo di esperti ad alto livello sull'IA».

[47] Si veda la relazione della Commissione, nell'ambito della comunicazione COM(2020) 65 final, cit., paragrafo 1.1. Sul punto, criticamente, G. Marchianò, cit., 8, secondo la quale «In verità, nella proposta non è dato rinvenire un quadro giuridico "solido" anche se, data la complessità della materia e il suo continuo divenire, risulta sicuramente difficile cogliere tale obiettivo».

[48] Sul punto, F. Pizzetti, *La protezione dei dati personali dalla direttiva al nuovo regolamento*, cit., 108. Sulle interrelazioni tra IA e protezione dei dati, si veda Memoria del Garante per la protezione dei dati personali – COM 2021(206) Proposta di regolamento (UE) sull'intelligenza artificiale, presentata il 9 marzo 2022 alle Commissioni IX e X riunite della Camera dei Deputati, in https://www.garanteprivacy.it/web/guest/home/docweb/-/docweb-display/docweb/9751565.

re del trattamento, supportato in determinati settori da certificazioni e codici di condotta[49], che contribuiscono a identificare i requisiti di conformità alle prescrizioni del GDPR.

In questa prospettiva, la proposta di regolamento sull'IA contiene, nell'allegato III, un elenco di sistemi di IA ad alto rischio, in considerazione di rischi che si sono già concretizzati o potrebbero verificarsi nel prossimo futuro. Tuttavia, al fine di garantire il necessario adeguamento alle nuove applicazioni dell'intelligenza artificiale, tale elenco può essere ampliato dalla Commissione, in determinati settori[50]. In particolare, l'art. 84 prevede un riesame di detto allegato una volta l'anno, dopo l'entrata in vigore del regolamento; mentre, ogni quattro anni la Commissione è tenuta a trasmettere al Parlamento europeo e al Consiglio una relazione di valutazione e riesame del regolamento, nonché una valutazione di impatto e di efficacia dei codici di condotta; e può inoltre proporre modifiche al regolamento stesso, tenuto conto del progresso tecnologico e dell'evoluzione della società dell'informazione.

Al fine di assicurare una più efficace tutela dei diritti e delle libertà delle persone all'interno del UE, le regole stabilite nella proposta di regolamento si applicano ai fornitori di sistemi IA, secondo un principio di non discriminazione, «a prescindere dal fatto che siano stabiliti nell'Unione o in un paese terzo, e agli utenti dei sistemi di IA stabiliti nell'Unione»[51], come precisato nell'art. 2, che definisce l'ambito di applicazione del regolamento[52].

Tuttavia, la proposta esclude dal proprio ambito di applicazione la cooperazione internazionale in materia di attività di contrasto[53], con una disposizione

---

[49] Cfr. *ibidem.*

[50] Cfr. relazione della Commissione, nell'ambito della comunicazione COM(2020) 65 final, cit., paragrafo 5.2.3.

[51] Considerando (10) della proposta di regolamento. In proposito, il successivo considerando (11) chiarisce che «è opportuno che determinati sistemi di IA rientrino nell'ambito di applicazione del presente regolamento anche quando non sono immessi sul mercato, né messi in servizio, né utilizzati nell'Unione. È il caso, ad esempio, di un operatore stabilito nell'Unione che appalta alcuni servizi a un operatore stabilito al di fuori dell'Unione in relazione a un'attività che deve essere svolta da un sistema di IA che sarebbe classificato ad alto rischio e i cui effetti avrebbero un impatto sulle persone fisiche che si trovano nell'Unione. In tali circostanze il sistema di IA utilizzato dall'operatore al di fuori dell'Unione potrebbe trattare dati raccolti nell'Unione e da lì trasferiti, nel rispetto della legge, e fornire all'operatore appaltante nell'Unione l'output di tale sistema di IA risultante da tale trattamento, senza che tale sistema di IA sia immesso sul mercato, messo in servizio o utilizzato nell'Unione. Al fine di impedire l'elusione del presente regolamento e di garantire una protezione efficace delle persone fisiche che si trovano nell'Unione, è opportuno che il presente regolamento si applichi anche ai fornitori e agli utenti di sistemi di IA stabiliti in un paese terzo, nella misura in cui l'output prodotto da tali sistemi è utilizzato nell'Unione».

[52] Precisamente, l'art 2 stabilisce che il regolamento si applica: «ai fornitori che immettono sul mercato o mettono in servizio sistemi di IA nell'Unione, indipendentemente dal fatto che siano stabiliti nell'Unione o in un paese terzo; b) agli utenti dei sistemi di IA situati nell'Unione; c) ai fornitori e agli utenti di sistemi di IA situati in un paese terzo, laddove l'output prodotto dal sistema sia utilizzato nell'Unione».

[53] Art. 2, comma 4.

che ha suscitato «gravi preoccupazioni nell'EDPB e nel GEPD, considerato che comporta un rischio significativo di elusione (ad esempio nel caso di paesi terzi o di organizzazioni internazionali che gestiscono applicazioni ad alto rischio su cui fanno affidamento le autorità pubbliche dell'UE)»[54].

Il *draft* del regolamento contiene, all'art 3, una serie di definizioni, tra cui quella di intelligenza artificiale, precisando che con tale termine si intende «un software sviluppato con una o più delle tecniche e degli approcci elencati nell›allegato I, che può, per una determinata serie di obiettivi definiti dall›uomo, generare output quali contenuti, previsioni, raccomandazioni o decisioni che influenzano gli ambienti con cui interagiscono». Nel quadro giuridico delineato, la definizione di IA mira ad essere «il più possibile neutrale» dal punto di vista tecnologico, anche in considerazione degli sviluppi futuri e dell'evoluzione del mercato digitale, come chiarito nella relazione di accompagnamento alla proposta stessa.

Lo stesso art. 3 definisce, inoltre, i partecipanti all'intera catena di valore dell'IA, quali i "fornitori" e gli "utenti" di sistemi IA, includendovi operatori sia pubblici che privati[55].

L'impianto normativo della proposta segue un approccio basato sul rischio, in modo da promuovere lo sviluppo dell'IA, garantendo un livello elevato di protezione degli interessi pubblici, quali la salute e la sicurezza, e dei diritti fondamentali (considerando 5).

Nell'articolato della proposta, tuttavia, come rilavato nel parere congiunto dell'EDPB e del GEPD, non sono presi specificamente in considerazione i rischi per categorie di persone o per la società nel suo complesso (ad esempio effetti collettivi di particolare rilevanza come la discriminazione di gruppo o la manifestazione di opinioni politiche in spazi accessibili al pubblico)[56].

---

[54] EDPB-GEPD, Parere congiunto 5/2021, sulla proposta di regolamento del Parlamento europeo e del Consiglio che stabilisce regole armonizzate sull'intelligenza artificiale (legge sull'intelligenza artificiale), 18 giugno 2021, 3 e 9.

[55] Sul punto, Camera dei deputati, Ufficio rapporti con l'Unione europea, Dossier n. 57 del 12 novembre 2021, in http://documenti.camera.it/leg18/dossier/pdf/ES057.pdf?_1647278357058, in cui si rileva che «Al riguardo, si ricorda che nella relazione trasmessa al Parlamento ai sensi dell'art. 6, comma 4, della legge n. 234 del 2012, il Governo, nell'ambito delle prospettive negoziali e delle eventuali modifiche ritenute necessarie od opportune, sottolinea in via generale l'elasticità del perimetro di applicazione del nuovo regime (in quanto modificabile attraverso atti delegati), e la necessità di valutare il rischio di incertezza giuridica e di "delega in bianco" alla Commissione».

[56] Ivi, 10, i quali raccomandano che «anche i rischi derivanti dall'IA per la società e le relative categorie siano valutati e attenuati. L'EDPB e il GEPD sono del parere che dovrebbe essere chiarito l'approccio basato sul rischio adottato nella proposta uniformando al GDPR il concetto di «rischio per i diritti fondamentali» per quanto riguarda gli aspetti correlati alla protezione dei dati personali. Che si tratti di utenti finali, semplici interessati o altre persone coinvolte dal sistema di IA, l'assenza nel testo della proposta di qualsiasi riferimento alla persona sulla quale si producono gli effetti del sistema di IA si configura come una grave carenza. Infatti, gli obblighi degli operatori nei confronti delle persone interessate dovrebbero emanare più concretamente dalla tutela del singolo e dei suoi diritti. Pertanto, l'EDPB e il GEPD sollecitano i legislatori ad affrontare esplicitamente nella proposta la questione dei diritti e dei mezzi di ricorso a disposizione delle persone soggette ai sistemi di IA».

I rischi intrinseci all'uso dell'IA, anche legati alla programmazione della macchina, possono produrre effetti discriminatori in grado di incidere su intere classi di persone, come si evince dall'esperienza canadese sopra ricordata.

Il processo decisionale automatizzato utilizza un set di dati, che rappresentano la base del processo di addestramento della macchina, e quindi la qualità degli *input* di formazione influisce sulla qualità degli *output* generati dal sistema[57]. Un esempio classico è costituito dal riconoscimento facciale, in quanto il *software* addestrato sui volti di persone bianche potrebbe essere meno capace di identificare accuratamente gli individui con carnagione più scura[58].

In sostanza, i valori, le assunzioni, i pregiudizi coinvolti nel set di dati di formazione, nonché la definizione dei dati ritenuti rilevanti, incide sui risultati, che ne riproducono carenze e parzialità[59].

I sistemi automatizzati possono quindi fare affidamento su variabili discriminatorie, spesso difficili da correggere e perfino da rilevare.

Da questa angolazione, emerge sia l'importanza di adottare impostazioni – il più possibile – obiettive, che favoriscano la generazione di "*output*" neutrali, tenuto conto che la scelta sugli elementi rilevanti possa avere una connotazione intrinsecamente politica[60], sia – a maggior ragione – l'esigenza di implementare in ogni caso test di verifica, in modo da controllare l'incidenza delle scelte operate e prevenire impatti pregiudizievoli.

Il *draft* di regolamento prevede una serie di pratiche vietate, puntualmente elencate nell'art. 5, in considerazione dei livelli di rischio legati all'applicazione di determinati sistemi di IA, quali l'uso di tecniche subliminali suscettibili di distorcere il comportamento delle persone o di sfruttare le vulnerabilità di gruppi sociali; ovvero, l'impiego da parte delle autorità pubbliche o per loro conto di sistemi che implichino l'attribuzione di un punteggio sociale, basato sul monitoraggio e la valutazione di comportamenti individuali; nonché il ricorso all'identificazione biometrica, in tempo reale, in spazi accessibili al pubblico a fini di attività di contrasto, salvo il verificarsi di condizioni di stretta necessità, specificamente indicate, per il perseguimento o la prevenzione di determinati reati[61].

---

[57] P. Molnar, L. Gill, *Bots at the gate, A human rights analysis of automated decision-making in Canada's immigration and refugee system,* cit., 9, dove, con riferimento al caso dell'applicazione di processi decisionali automatizzati ai diritti di immigrazione dei rifugiati in Canada, si rileva che «As a recent report from the Canadian company Integrate.ai points out, "[I]n standard practice, machine learning assumes the future will look like the past. When the past is unfair or biased, machine learning will propagate these biases and enhance them through feedback loops"».

[58] *Ibidem*.

[59] *Ibidem.*

[60] Ivi, 10.

[61] Ai sensi dell'art. 5, comma 1, lett. *d*), della proposta di regolamento è vietato «l'uso di sistemi di identificazione biometrica remota "in tempo reale" in spazi accessibili al pubblico a fini di attività di contrasto, a meno che e nella misura in cui tale uso sia strettamente necessario per uno dei seguenti obiettivi: i) la ricerca

Il ricorso a tale sistema è subordinato a un'autorizzazione preventiva rilasciata da un'autorità giudiziaria o da un'autorità amministrativa; tuttavia, in una «situazione di urgenza debitamente giustificata, è possibile iniziare a usare il sistema senza autorizzazione e richiedere l'autorizzazione solo durante o dopo l'uso» (art. 5, comma 3).

Con riferimento all'identificazione biometrica da remoto «in tempo reale» in spazi accessibili al pubblico, l'EDPB e il GEPD, nel parere congiunto già ricordato, hanno ritenuto che l'approccio adottato nella proposta di regolamento «sia viziato sotto numerosi aspetti», evidenziando l'invasività del trattamento, anche a prescindere dalla circostanza che l'identificazione sia eseguita in tempo reale, e la capacità dei sistemi di identificazione di massa di identificare migliaia di persone in pochissime ore. L'uso di tale tecnologia da remoto *a posteriori* nel contesto di una manifestazione di protesta politica può avere un pesante effetto dissuasivo sull'esercizio dei diritti e delle libertà fondamentali, come la libertà di riunione e di associazione e, più in generale, sui principi fondanti della democrazia[62]. La proposta «sembra non tenere conto del fatto che, nel monitoraggio di aree all'aperto, gli obblighi previsti dalla legislazione dell'UE in materia di protezione dei dati devono essere adempiuti in riferimento non soltanto alle persone sospette, ma a tutte quelle di fatto monitorate»[63].

Pertanto, l'EDPB e il GEPD hanno sollecitato l'inclusione di «un divieto generale di qualsiasi uso dell'IA a fini di riconoscimento automatico, in spazi accessibili al pubblico, delle caratteristiche umane – come il volto ma anche l'andatura, le impronte digitali, il DNA, la voce, le sequenze di battute su tastiera e altri segnali biometrici o comportamentali – in qualsiasi contesto».

Come osservato dal Garante per la protezione dei dati personali, la tassonomia dei divieti rievoca i parametri previsti dal GDPR, e relativamente alla rilevazione biometrica a fini di pubblica sicurezza (oltre che, a fortiori, a meri fini facilitativi nel settore privato) positivizza una prassi ormai consolidata delle Autorità

---

mirata di potenziali vittime specifiche di reato, compresi i minori scomparsi; ii) la prevenzione di una minaccia specifica, sostanziale e imminente per la vita o l'incolumità fisica delle persone fisiche o di un attacco terroristico; iii) il rilevamento, la localizzazione, l'identificazione o l'azione penale nei confronti di un autore o un sospettato di un reato di cui all'articolo 2, paragrafo 2, della decisione quadro 2002/584/GAI del Consiglio, punibile nello Stato membro interessato con una pena o una misura di sicurezza privativa della libertà della durata massima di almeno tre anni, come stabilito dalla legge di tale Stato membro». Il comma 2 precisa che l'uso di tali strumenti dovrà tener conto della natura della situazione e delle conseguenze sui diritti e le libertà delle persone, nel rispetto di condizioni di necessità e proporzionalità, in particolare per quanto riguarda le limitazioni temporali, geografiche e personali.

[62] EDPB-GEPD, cit., 13. Sullo stesso aspetto della disciplina cfr. anche M. Veale, F. Zuiderveen Borgesius, *Demystifying the Draft EU Artificial Intelligence Act. Analysing the good, the bad, and the unclear elements of the proposed approach*, in *Computer Law Rev Int*, 2021, 4, 101, I. Barkane, *Intelligence Act: The need for prohibitions and a stricter approach to biometric surveillance*, in *Information Polity*, 2022, 27, 147 ss., oltre alla dottrina citata nella nota 9 *infra*.

[63] *Ibidem*.

di protezione dati, che ne escludono l'ammissibilità, «in assenza di previsioni normative corredate di idonee garanzie», secondo un indirizzo confermato dal legislatore in sede di conversione del d.l. n. 139/2021[64].

La stessa Autorità suggerisce in ogni caso un approccio più garantista, che vieti il ricorso a sistemi di IA funzionali all'attribuzione di punteggi sociali, alla deduzione di emozioni o alla classificazione di gruppi sociali, sulla base di dati biometrici o di fattori discriminanti[65].

I sistemi di IA che categorizzano le persone, riducendole ad insiemi, attraverso l'uso di dati biometrici, possono infatti incidere in modo determinante sulla libertà delle persone e sulla dignità umana, come si dirà oltre.

4. *I sistemi di intelligenza artificiale ad alto rischio*

La struttura del *draft* di regolamento è incardinata sul rischio e sulla relativa graduazione.

L'art. 6 individua una serie di sistemi di IA definiti ad alto rischio, che sono assoggettati ad una procedura di valutazione di conformità *ex ante*[66], volti a garantire che funzionino in modo coerente con le finalità previste e rispettino i requisiti stabiliti nel regolamento (art. 19 e 43); e sistemi di IA a rischio limitato, per i quali sono previsti specifici obblighi di trasparenza (art. 52)[67].

---

[64] Memoria del Garante per la protezione dei dati personali – COM 2021 (206) Proposta di regolamento (UE) sull'intelligenza artificiale, cit.

[65] *Ibidem.* Analogamente, EDPB-GEPD, cit., 14.

[66] In proposito, EDPB-GEPD, cit., 14 ss., «valutano positivamente il fatto che i sistemi di IA che comportano rischi elevati debbano essere sottoposti a una preventiva valutazione della conformità prima di poter essere immessi sul mercato o altrimenti resi operativi nell'UE. In linea di principio, questo modello normativo è benaccetto in quanto costituisce un buon punto di equilibrio tra la disponibilità all'innovazione e un elevato livello di protezione proattiva dei diritti fondamentali. Per poterlo applicare in ambienti specifici, come i processi decisionali delle istituzioni dei servizi pubblici o in infrastrutture essenziali, è necessario stabilire le modalità con cui esaminare il codice sorgente completo»; e «invitano ad adattare la procedura di valutazione della conformità di cui all'articolo 43 della proposta in modo tale da prevedere l'obbligo generale di sottoporre i sistemi di IA ad alto rischio a una valutazione della conformità ex ante da parte di terzi. Benché una valutazione della conformità da parte di terzi dei trattamenti ad alto rischio di dati personali non costituisca un requisito a norma dell'RGPD o dell'EUDPR, i rischi posti dai sistemi di IA devono ancora essere compresi nella loro interezza. L'inserimento di una previsione generale di condurre una valutazione obbligatoria della conformità a opera di terzi permetterebbe, quindi, di accrescere ulteriormente la certezza del diritto e la fiducia in tutti i sistemi di IA ad alto rischio».

[67] L'art. 52 della proposta di regolamento prevede che «I fornitori garantiscono che i sistemi di IA destinati a interagire con le persone fisiche siano progettati e sviluppati in modo tale che le persone fisiche siano informate del fatto di stare interagendo con un sistema di IA, a meno che ciò non risulti evidente dalle circostanze e dal contesto di utilizzo. Tale obbligo non si applica ai sistemi di IA autorizzati dalla legge per accertare, prevenire, indagare e perseguire reati, a meno che tali sistemi non siano a disposizione del pubblico per segnalare un reato» (comma 1). Lo stesso articolo stabilisce una serie di obblighi di trasparenza, con riferimento a un sistema di riconoscimento delle emozioni o un sistema di categorizzazione biometrica, assicurando che le per-

Analogamente al Libro bianco, il *draft* di regolamento impone controlli sulle applicazioni ad "alto rischio", ma stabilisce un quadro di conformità dei sistemi IA focalizzato sull'idoneità a generare un alto rischio sulla salute, la sicurezza e i diritti fondamentali, mentre il Libro bianco si basa essenzialmente su una classificazione secondo il settore di applicazione e l'uso previsto.

I sistemi a elevato rischio ricomprendono due categorie principali, ovvero sistemi di IA destinati a essere utilizzati come componenti di sicurezza di un prodotto, soggetti a valutazione di conformità *ex ante* da parte di terzi; e sistemi indipendenti, elencati nell'allegato III, suscettibili di incidere sensibilmente sui diritti fondamentali delle persone.

La Commissione può, tuttavia, ampliare tale elenco, secondo criteri e metodologie di valutazione del rischio indicate nell'art. 7 della proposta di regolamento.

La disposizione di cui all'art. 9 delinea un sistema di gestione del rischio come un processo in divenire, che richiede costante aggiornamento nell'intero ciclo di vita del sistema di IA.

Tale processo include: *(i)* l'identificazione e analisi dei rischi noti e prevedibili associati a ciascun sistema di IA ad alto rischio; *ii)* la stima e valutazione dei rischi che possono emergere quando il sistema di IA ad alto rischio è usato conformemente alla sua finalità prevista e in condizioni di uso improprio ragionevolmente prevedibile; *iii*) la valutazione di altri eventuali rischi derivanti dall'analisi dei dati raccolti dal sistema di monitoraggio successivo all'immissione sul mercato; *(iv)* l'adozione di adeguate misure di gestione dei rischi.

La norma mira a valutare i rischi residui associati a ciascun pericolo nonché il rischio residuo complessivo dei sistemi di IA ad alto rischio considerati accettabili, in caso di uso conforme alla finalità prevista, o di uso improprio ragionevolmente prevedibile. I rischi residui devono essere comunicati all'utente (art. 9, comma 4).

In ogni caso, per i sistemi ad alto rischio, l'art 14 prevede l'intervento umano, qualora si riscontrino distorsioni rispetto agli obiettivi dell'IA[68].

La proposta di regolamento stabilisce, inoltre, i requisiti in materia di dati e *governance* dei dati, prevedendo in particolare che i set di dati di addestramento, convalida e prova debbano essere «pertinenti, rappresentativi, esenti da errori e completi» (art. 10) e sancisce una serie di obblighi, tra cui quello di redigere la documentazione tecnica di un sistema di IA ad alto rischio prima dell'immis-

sone fisiche esposte siano informate del funzionamento del sistema; nonché in relazione ai sistemi di IA che generano o manipolano immagini o contenuti audio o video che assomigliano notevolmente a persone, oggetti, luoghi o altre entità o eventi esistenti e che potrebbero apparire falsamente autentici o veritieri per una persona ("*deep fake*"), in modo da chiarire che il contenuto è stato generato artificialmente.

[68] Sul punto, G. Marchianò, cit., 13, in cui si osserva che «In verità nonostante l'adozione di tale criterio, non sempre nella proposta di regolamento è dato intravedere un'organica distinzione tra IA ad alto rischio e IA a rischio minimo: ciò lo si evince soprattutto negli allegati che accompagnano gli enunciati dell'art. 6, rispetto all'art. 7 che dovrebbero riguardare i sistemi d'IA ad alto rischio».

sione sul mercato o della messa in servizio del sistema e di provvedere al relativo aggiornamento (art. 11); o di assicurare che i sistemi siano progettati in modo da consentire la registrazione automatica degli eventi ("log") durante il loro funzionamento, con un approccio complessivo che mira a bilanciare lo sviluppo tecnologico e la tutela dei diritti.

In concreto, la verifica che i set di dati siano «esenti da errori e completi» presenta diversi profili di criticità.

Come già osservato, con riferimento al caso dei sistemi di analisi predittiva implementati in Canada al fine di automatizzare determinate attività amministrative nel settore dell'immigrazione, il set di dati utilizzato per la formazione della macchina può contenere una serie di carenze e parzialità, difficili da prevedere e identificare. Nell'esempio, sinteticamente richiamato, non è stato affatto semplice per i ricercatori risalire alle carenze del set di addestramento che si riflettevano, com'è logico, sugli *output* del sistema.

Tale caso non è né isolato né raro, ma paradigmatico di un rischio immanente nel processo di *training*, derivante dalla molteplicità delle variabili che compongono il *data set* e dalla obiettiva difficoltà, spesso impossibilità, di prevedere possibili errori e incompletezze.

Tenendo conto degli impatti connessi all'uso di sistemi di IA altamente complessi, e della relativa incidenza sui diritti e le libertà delle persone, la verifica sul set di *training*, pur prevista dall'art. 10 del *draft* di regolamento[69], sarebbe di per sé insufficiente, se non associata ad un controllo *ex post* sugli *output* derivanti dall'implementazione del sistema di IA, quale controprova della coerenza del sistema con le finalità ad esso sottese[70]. Tale verifica appare fondamentale al fine di correggere le distorsioni e rimodulare, ove necessario, il *data set* di addestramento.

Sotto questo profilo, il monitoraggio continuo del sistema e la realizzazione di *audit* svolgono un ruolo cruciale per consentire uno sviluppo sicuro dell'IA e costruire un ecosistema di fiducia.

Il *draft* di regolamento prevede che prima dell'avvio di un sistema IA ad alto rischio e della sua immissione su mercato sia redatta una documentazione tecnica (art. 11, paragrafo 1), da aggiornare regolarmente, che dimostri la conformità

[69] L'art. 10 include nel sistema di *governance* dei dati «una valutazione preliminare della disponibilità, della quantità e dell'adeguatezza dei set di dati necessari; un esame atto a valutare le possibili distorsioni; l'individuazione di eventuali lacune o carenze nei dati e il modo in cui tali lacune e carenze possono essere colmate» (comma 2, lett. e) f), g)) e consente di trattare, in caso di stretta necessità, alcuni dati personali e particolari, al fine di garantire il monitoraggio.

[70] Tale processo di verifica *ex post* si potrebbe svolgere, ad esempio, anche con il supporto dell'IA, con successivo controllo umano, attraverso algoritmi di controllo che utilizzino un set diverso di dati, in funzione di verifica, in modo da controllare che gli *output* siano coerenti con le finalità sottese all'uso dell'IA in un determinato contesto.

ai requisiti relativi all'alto rischio e contenga, come minimo, le informazioni di cui all'allegato IV[71].

Nella fase di monitoraggio, successiva all'immissione sul mercato, funzionale alla valutazione di «altri eventuali rischi» connessi all'uso dei sistemi IA[72], la proposta di regolamento prevede la predisposizione di un piano di monitoraggio, facente parte della stessa documentazione tecnica di cui all'allegato IV, da redigere secondo disposizioni dettagliate da adottare attraverso atti esecutivi della Commissione, chiamata a definire il modello del piano e gli elementi che lo compongono.

In tale fase, lo sviluppo di una metodologia di verifica costante degli *output*, che implichi l'eventuale aggiornamento del set di *training,* rappresenta un contrappeso essenziale per assicurare il rispetto dei requisiti stabiliti dal regolamento anche in materia di *governance* dei dati.

In ogni caso, i sistemi di IA ad alto rischio devono essere sottoposti ad una valutazione di conformità (art. 19), secondo le procedure indicate nell'art. 43.

---

[71] La documentazione tecnica di cui all'allegato IV, paragrafo 2 include, tra l'altro, una descrizione dettagliata degli elementi del sistema di IA e del processo relativo al suo sviluppo, che comprende «a) i metodi applicati e le azioni eseguite per lo sviluppo del sistema di IA, compresi, ove opportuno, il ricorso a sistemi o strumenti preaddestrati forniti da terzi e il modo in cui sono stati utilizzati, integrati o modificati dal fornitore; b) le specifiche di progettazione del sistema, vale a dire la logica generale del sistema di IA e degli algoritmi; le principali scelte di progettazione, comprese le motivazioni e le ipotesi formulate, anche per quanto riguarda le persone o i gruppi di persone sui quali il sistema è destinato a essere utilizzato; le principali scelte di classificazione; gli aspetti che il sistema è progettato per ottimizzare e la pertinenza dei diversi parametri; le decisioni in merito a eventuali compromessi posti in essere con riguardo alle soluzioni tecniche adottate per soddisfare i requisiti di cui al titolo III, capo 2; c) la descrizione dell'architettura del sistema che spiega in che modo i componenti software si basano l'uno sull'altro o si alimentano reciprocamente e si integrano nel processo complessivo; le risorse computazionali utilizzate per sviluppare, addestrare, sottoporre a prova e convalidare il sistema di IA; d) ove pertinente, i requisiti in materia di dati mediante schede tecniche che descrivono le metodologie e le tecniche di addestramento e i set di dati di addestramento utilizzati, comprese le informazioni sull'origine di tali set di dati, sul loro ambito di applicazione e sulle loro principali caratteristiche; le modalità di ottenimento e di selezione dei dati; le procedure di etichettatura, ad esempio per l'apprendimento supervisionato, e le metodologie di pulizia dei dati, ad esempio il rilevamento di valori anomali (outlier); e) la valutazione delle misure di sorveglianza umana necessarie in conformità dell'articolo 14, compresa una valutazione delle misure tecniche necessarie per facilitare l'interpretazione degli output dei sistemi di IA da parte degli utenti, in conformità dell'articolo 13, paragrafo 3, lettera d); f) ove applicabile, una descrizione dettagliata delle modifiche predeterminate del sistema di IA e delle sue prestazioni, unitamente a tutte le informazioni pertinenti relative alle soluzioni tecniche adottate per garantire la conformità costante del sistema di IA ai requisiti pertinenti di cui al titolo III, capo 2; g) le procedure di convalida e di prova utilizzate, comprese le informazioni sui dati di convalida e di prova utilizzati e sulle loro principali caratteristiche; le metriche utilizzate per misurare l'accuratezza, la robustezza, la cibersicurezza e la conformità ad altri requisiti pertinenti di cui al titolo III, capo 2, nonché gli impatti potenzialmente discriminatori; i log delle prove e tutte le relazioni di prova corredate di data e firma delle persone responsabili, anche per quanto riguarda le modifiche predeterminate di cui alla lettera f)».

[72] L'art. 9 del draft di regolamento istituisce un sistema di gestione dei rischi che include, tra l'altro, la valutazione di «altri eventuali rischi derivanti dall'analisi dei dati raccolti dal sistema di monitoraggio successivo all'immissione sul mercato di cui all'articolo 61». Tale ultima disposizione prevede un sistema di monitoraggio «proporzionato alla natura delle tecnologie di intelligenza artificiale e ai rischi del sistema di IA ad alto rischio».

Tale disposizione distingue una procedura di valutazione basata sul controllo interno, disciplinata dall'allegato VI, rimessa allo stesso fornitore che ha il compito di valutare autonomamente la conformità del proprio sistema di gestione della qualità e della documentazione tecnica ai requisiti normativi; e una procedura che prevede l'intervento di soggetti terzi, disciplinata dall'allegato VII.

La previsione che affida all'autovalutazione molteplici sistemi di IA non riflette tuttavia l'esigenza di garantire la protezione di diritti fondamentali delle persone, a fronte di applicazioni tecnologiche ad alto rischio, spesso opache e imprevedibili, né contribuisce a sviluppare quel clima di fiducia che pur rappresenta uno degli obiettivi fondamentali della stessa proposta di regolamento[73].

Gli stessi fornitori si troverebbero ad affrontare una sfida etica, nel bilanciamento tra i vantaggi dell'innovazione e della immissione sul mercato delle tecnologie dagli stessi prodotti e la responsabilità sociale, in assenza di una valutazione da parte di soggetti terzi[74].

Sotto questo profilo, la previsione di una supervisione obiettiva obbligatoria per tutti i sistemi di IA potrebbe favorire lo sviluppo di un'intelligenza artificiale sicura ed etica, in un quadro normativo affidabile.

5. *Quali dati? Tipologia di dati e algoritmi intelligenti*

Come già rilevato nei precedenti paragrafi, lo sviluppo degli ultimi anni di tecnologie di intelligenza artificiale è stato permesso innanzitutto da una disponibilità di grandi quantità di dati (*big data*), che possono essere archiviati, analizzati ed elaborati grazie all'aumento della capacità computazionale[75]. L'enorme quan-

[73] Cfr. J. MöKander, M. Axente, F. Casolari, L. Floridi, *Conformity Assessments and Post-market Monitoring: a Guide to the Role of Auditing in the Proposed European AI Regulation,* in *Minds and Machines*, 32, 2022, 241 ss., partic., 262, in cui si evidenzia che «Checks and balances. Although high-risk AI systems are subject to conformity assessments, the enforcement of the requirements set out in the AIA is less stringent than it appears (MacCarthy & Propp, 2021). This is primarily because – for most high-risk AI systems – the conformity assessments will be based on internal checks conducted by the system provider itself. Similarly, while providers must draw up an EU declaration of conformity and give a copy of it to the relevant national authorities upon request (AIA ["Artificial Intelligence Act"], Article 48), how providers ensure compliance with the AIA is not disclosed to the public. This lack of checks and balances may be problematic because pursuing rapid technological progress leaves little time to ensure that AI systems are robust and ethical (Whittlestone et al., 2019b)»; EDPB-GEPD, cit., 14 ss.; G. Contissa, F. Galli, F. Godano, G. Sartor, cit., 30.

[74] Cfr. J. MöKander, M. Axente, F. Casolari, L. Floridi, cit., 262, in cui si rileva che « Moreover, there is always a risk of adversarial behaviour during conformity assessments. Thus, companies find themselves wedged between the benefits of innovation and social responsibility and may not act ethically in the absence of oversight (Turner Lee, 2018)».

[75] Si veda, oltre a quanto indicato nella nota 3, anche D.E. Holmes, *Big Data. A very short introduction*, Oxford, 2017, 14 ss. M. Delmastro, A. Nicita, *Big Data. Come stanno cambiano il nostro mondo*, Bologna, 2019, 10 ss.; L. Piatti, *Big Data*, in G. Ziccardi, P. Perri (a cura di), *Dizionario Legal Tech*, Milano, 2021, 102 ss. V.

tità di dati (considerata in grandi insiemi definiti *data lake*) può essere analizzata in modo utile e affidabile grazie a macchine con una capacità di calcolo estremamente elevata e tale da poter estrarre informazioni e ricavarne elaborazioni "intelligenti", sulla base degli algoritmi utilizzati.

Le componenti fondamentali dello sviluppo di tecnologie di intelligenza artificiale sono dunque i dati e gli algoritmi, come evidenziato dalla stessa Commissione europea nel processo che ha portato alla proposta di regolamento sull'IA[76].

Le caratteristiche quantitative, ma anche qualitative, dei dati sono dunque un presupposto fondamentale, tanto quanto le caratteristiche dell'algoritmo, per tecnologie di intelligenza artificiale efficaci. Come dimostra il noto progetto "Norman" del *Massachusetts Institute of Technology*, lo stesso algoritmo può portare a risultati (molto) diversi se "istruito" su *dataset* diversi[77].

I dati utilizzati dalle tecnologie di intelligenza artificiale possono essere, da un punto di vista informatico, di tipo strutturato (ordinato in *database* relazionali facilmente elaborabili), non strutturato (qualsiasi dato ricavato senza classificazioni da file di testo, audio, foto, video, messaggi etc.) o semi-strutturato (dati non strutturati ma con alcuni metadati che possono essere utilizzati per la loro analisi ed elaborazione)[78].

Dal punto di vista giuridico, la classificazione più rilevante è invece quella che distingue i dati "personali" da quelli "non personali", sulla base della definizione fondamentale contenuta nel GDPR[79]. In realtà, le tecnologie di intelligenza artificiale e i *big data* in qualche modo mettono in crisi la rilevanza di tale

anche AGCM, AGCOM, Garante per la protezione dei dati personali, *Big Data. Indagine conoscitiva congiunta. Rapporto finale,* 2020, in part. 7-8.

[76] Cfr. il già citato Libro Bianco sull'intelligenza artificiale: Commissione europea, 19.2.2020, COM(2020) 65 final, cit., 18.

[77] *Norman*, così nominato facendo riferimento all'assassino schizofrenico della pietra miliare di Alfred Hitchcok *Psycho*, viene definito come la «prima intelligenza artificiale psicopatica». In sintesi, il gruppo di ricerca del *MIT* nel corso del 2018 ha sottoposto un algoritmo di *machine learning* pensato per descrivere le immagini con testo ad un processo di apprendimento supervisionato basato su un *data set* di immagini inquietanti. Per un altro algoritmo con caratteristiche simili sono state usate immagini comuni (prese dalla banca dati di Microsoft MSCOCO). All'esito del processo di apprendimento ad entrambi gli algoritmi sono state proposte le immagini del test di Rorschach, cui *Norman* ha risposto con descrizioni di scene di morte, mentre l'altro algoritmo con descrizioni più positive. La sintesi e i risultati del progetto sono disponibili nel sito *http://norman-ai.mit.edu/*.

[78] D.E. Holmes, cit., 5 ss., in cui si legge che l'80% dei dati sono in formato non strutturato.

[79] Ai sensi dell'art. 4, n. 1) del GDPR è 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». Come noto, la definizione è presa come riferimento anche per l'ambito di applicazione del regolamento sulla libera circolazione dei dati non personali (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),

distinzione fondamentale, da almeno due prospettive. Da un lato, infatti, il volume e la varietà dei dati personali rendono più frequente il ricorso a insiemi di dati "misti"[80] difficilmente separabili specie se non strutturati. I *database* misti includono sia dati personali che dati non personali, ma per questi ultimi, se «indissolubilmente legati» ai primi, devono trovare applicazione le regole relative ai dati personali[81]. Dall'altro lato, l'impiego di tecnologie IA può incidere sulla qualificazione in concreto di un determinato dato come "personale" o meno: gli algoritmi intelligenti possono infatti creare connessioni tra dati e completare le informazioni disponibili su dati apparentemente non personali o anonimizzati, estendendo così drasticamente l'ambito di "identificabilità" indiretta di una persona fisica e, quindi, l'ambito di applicazione della definizione di dato personale[82]. L'orientamento della dottrina che evidenzia la crisi della distinzione tra dati personali e non personali si basa anche su questi argomenti[83].

Per ragioni analoghe, assume elementi di incertezza anche l'individuazione di «categorie particolari di dati», ambito entro cui, com'è noto, il GDPR raggruppa i dati personali maggiormente sensibili[84]: quanto visto in termini di consistenza dei *big data* e possibilità di analisi ed elaborazione sugli insiemi di dati da parte di algoritmi vale anche per tale distinzione[85].

---

che disciplina i dati «diversi dai dati personali definiti all'articolo 4, punto 1, del regolamento (UE) 2016/679» (così l'art. 3, lett. a) del regolamento).

[80] Così definiti nelle linee guida della Commissione UE sul regolamento sulla libera circolazione dei dati non personali – Comunicazione della Commissione europea del 29.5.2019, COM(2019) 250, par. 2.2.

[81] Art. 2, par. 2, regolamento sulla libera circolazione dei dati non personali, come interpretato anche dalle linee guida della Commissione citate. La Commissione precisa anche che, «ai fini pratici», il concetto di «indissolubilmente legati» riguarda sia il caso in cui la separazione dei dati personali e non personali è impossibile sia il caso in cui la separazione è ritenuta dal titolare del trattamento «economicamente inefficiente o non tecnicamente realizzabile» (Comunicazione della Commissione europea COM(2019) 250, cit.).

[82] In E.M. Weitzenboeck, P. Lison, M. Cyndecka, M. Langford, *The GDPR and unstructured data: is anonymization possible?*, in *Int. Data Privacy Law*, 3, 2022, 203 si fa l'esempio di tecnologie che permettono di riconoscere l'autore di un testo, individuare il soggetto di un'immagine indistinguibile dall'occhio umano o attribuire una voce contraffatta ad una persona. Sull'argomento cfr. anche M. Finck, F. Pallas, *They who must not be identified – Distinguishing personal from non-personal data under the GDPR*, in *Int. Data Privacy Law*, 1, 2020, 11.; I. Graef, R. Gellert, M. Husovec, *Towards a holistic regulatory approach for the European data economy: why the illusive notion of non-personal data is counterproductive to data innovation*, in *European Law Review*, 5, 2018, 605 ss. Evidenzia il problema anche la stessa Commissione europea nel Libro Bianco sull'intelligenza artificiale (Comunicazione della Commissione europea, 19.2.2020, COM(2020) 65 final, cit., 11).

[83] Cfr. in particolare N. Purtova, *The law of everything. Broad concept of personal data and future of EU data protection law*, in *Law, Inn. and Tech.*, 2018, 1, 41 ss., I. Graef, R. Gellert, M. Husovec, cit., 605 ss. V. anche G. De Gregorio, *Digital Constitutionalism in Europe. Reframing Rights and Powers in the Algorithmic Society*, Cambridge, 2022, 236 ss.

[84] L'articolo 9, par. 1, del GDPR include nelle «categorie particolari» i dati «che rivelino l'origine razziale o etnica, le opinioni politiche, le convinzioni religiose o filosofiche, o l'appartenenza sindacale, nonché trattare dati genetici, dati biometrici intesi a identificare in modo univoco una persona fisica, dati relativi alla salute o alla vita sessuale o all'orientamento sessuale della persona».

[85] Si veda sul tema specifico D. Clifford, M. Richardson, N. Witzleb, *Artificial Intelligence and Sensitive Inferences: New Challenges for Data Protection*, in J. Ford, J. Seoh, D. Thampapillai (a cura di), *Regulatory Insights*

6. *La "neutralità dei dati" nella proposta di regolamento e l'eccezione dei dati biometrici*

La proposta di regolamento sull'IA si inserisce nel contesto descritto considerando in linea di principio dati di qualsiasi natura e tipologia. Come esposto *supra* al paragrafo 4, lo schema di regolamento impone standard (piuttosto alti) ai dati di addestramento, convalida e prova nell'ambito dei sistemi ad alto rischio, senza però qualificare ulteriormente i dati presi in considerazione. La proposta è dunque di per sé "neutrale" rispetto ai dati regolati, considerando tecnologie che possono impiegare dati personali e non personali. Coerentemente con questo approccio, il regolamento sull'IA si propone di non interferire con la normativa europea in materia di dati personali. Nella relazione di accompagnamento della proposta si esplicita che questa non pregiudica il GDPR e la direttiva sulla protezione dei dati nelle attività di polizia e giudiziarie (direttiva (UE) 2016/680 – cd. "LED")[86]. Le previsioni del nuovo regolamento devono dunque essere coerenti con il contesto esistente del diritto europeo della protezione dei dati personali e con le garanzie ad esso connesse, che continuano ad applicarsi anche con il rango di diritti fondamentali in base al dettato della Carta di Nizza (articolo 8) e della giurisprudenza della Corte di giustizia[87].

Un'eccezione rilevante alla "neutralità" del nuovo atto, rispetto alla qualificazione dei dati impiegati nei sistemi di IA regolati, riguarda i dati biometrici. I sistemi di IA basati sulla biometrica sono particolarmente presenti all'interno della proposta di regolamento, che si sofferma su di essi fin dalle definizioni. Le tecnologie biometriche possono essere considerate «il canarino nella miniera» dell'intelligenza artificiale[88], ossia come il primo ambito in cui i rischi per i diritti legati all'impiego di tecnologie di IA possono manifestarsi[89].

La proposta definisce e disciplina in particolare tre sistemi di IA che si basano su dati biometrici: il sistema di «identificazione biometrica remota», il sistema di «riconoscimento delle emozioni» e il sistema di «categorizzazione biometrica».

---

on Artificial Intelligence, Cheltenham, 2022, 19 ss.

[86] Comunicazione della Commissione europea COM(2021) 206, cit., par. 1.2.

[87] Si veda sul punto, in ampia letteratura, F. Rossi Dal Pozzo, *La tutela dei dati personali nella giurisprudenza della Corte di giustizia*, in Aa.Vv. *Annali Aiusde*, Bari, 2020, 70 ss., e la dottrina ivi citata.

[88] L'espresssione è di C. Kind, *Containing the canary in the AI coalmine – the EU's efforts to regulate biometrics*, disponibile su *https://www.adalovelaceinstitute.org/blog*, 2021.

[89] Nell'ambito delle tecnologie biometriche un'attenzione maggiore da parte del pubblico è stata dedicata al riconoscimento facciale, con iniziative di gruppi di attivisti quali *Reclaim Your Face* (su cui si veda *https://reclaimyourface.eu/*) che hanno portato alla risoluzione del Parlamento UE sulla moratoria delle tecnologie di riconoscimento facciale (Risoluzione del Parlamento europeo del 6 ottobre 2021 sull'intelligenza artificiale nel diritto penale e il suo utilizzo da parte delle autorità di polizia e giudiziarie in ambito penale (2020/2016(INI)) e il divieto a livello nazionale con il già citato d.l. n. 139/2021.

Il sistema di identificazione biometrica remota viene definito quale sistema finalizzato all'identificazione a distanza di persone fisiche mediante il confronto dei dati biometrici di una persona con i dati biometrici contenuti in una banca dati di riferimento[90]. Come evidenziato sopra, tale sistema è inserito tra quelli vietati – e comunque con alcune eccezioni – soltanto nel caso di identificazione biometrica remota «in tempo reale» (ossia in cui il rilevamento dei dati biometrici, il confronto e l'identificazione avvengono contestualmente[91]) in «spazi accessibili al pubblico» e a fini di attività di contrasto[92]. I sistemi di identificazione biometrica remota qualificabili soltanto come «in tempo reale» e «a posteriori» (non in tempo reale) sono invece inseriti tra i sistemi di IA ad alto rischio di cui all'allegato III del regolamento. Si è già detto *supra* di come tali limitazioni possano essere problematiche, come rilevato da EDPB e GEPD nel parere sulla proposta di regolamento.

La seconda tecnologia basata sui dati biometrici, considerata espressamente dalla proposta di regolamento, è il «sistema di riconoscimento delle emozioni», definito come un sistema di IA «finalizzato all'identificazione o alla deduzione di emozioni o intenzioni di persone fisiche sulla base dei loro dati biometrici»[93].

L'ultima tecnologia considerata è il «sistema di categorizzazione biometrica», definito come un sistema di IA che utilizza i dati biometrici di persone fisiche al fine di assegnarle a categorie specifiche[94]. Per questo sistema e per quello di riconoscimento delle emozioni, la proposta di regolamento prevede un'informativa alle persone fisiche esposte al sistema in merito al suo funzionamento[95].

In tutte e tre le definizioni richiamate il fondamento è la nozione di "dati biometrici": quest'ultima categoria viene definita a sua volta dalla proposta di regolamento, riprendendo la definizione contenuta nel GDPR e nella LED. In base a tale definizione, sono dati biometrici i «dati personali ottenuti da un trattamento tecnico specifico relativi alle caratteristiche fisiche, fisiologiche o compor-

---

[90] Art. 3, n. 36) proposta di regolamento sull'AI.

[91] Art. 3, n. 37).

[92] Art. 5, par. 1, lett. d proposta di regolamento. Con riguardo a questo specifico aspetto, la base giuridica del regolamento proposto viene individuata nell'articolo 16 del Trattato sul Funzionamento dell'Unione europea, riferimento principale all'interno del Trattato per la protezione dei dati personali e base giuridica anche del GDPR. Per il resto della disciplina, la base giuridica della proposta di regolamento sull'intelligenza artificiale è l'articolo 114 del TFUE – ravvicinamento delle legislazioni per gli obiettivi del mercato interno; è significativo notare che tutte le altre proposte della Commissione relative alla strategia europea dei dati (regolamento sull'accesso equo ai dati e sul loro utilizzo – cd. *Data Act*, regolamento sulla governance dei dati – cd. *Data Governance Act*) e sui servizi digitali (regolamenti sul mercato unico dei servizi digitali – *Data Services Act* e sui mercati digitali – *Digital Markets Act*) indicano solo l'articolo 114 del TFUE come base giuridica.

[93] Art. 3, n. 34) proposta di regolamento sull'AI.

[94] Art. 3, n. 35).

[95] Art. 52, par. 2.

tamentali di una persona fisica che ne consentono o confermano l'identificazione univoca, quali l'immagine facciale o i dati dattiloscopici»[96].

Le tecnologie biometriche incluse nella disciplina del regolamento sull'IA sono dunque esclusivamente quelle che non solo coinvolgono dati personali, ma sono anche connotate dalla specifica nozione di dati biometrici propria del diritto europeo dei dati personali. Quest'ultima considera, come si nota dalla disposizione richiamata, soltanto i dati personali frutto di un «trattamento tecnico specifico» che «ne consentono o confermano l'identificazione univoca»[97]. Sulla base di queste limitazioni, ad esempio, una banca dati che contiene fotografie di volti non corrisponde alla nozione di dati biometrici in assenza di trattamenti tecnici specifici che consentano o confermino l'identificazione di persone fisiche[98].

Proprio tali limitazioni sono considerate fonte di ambiguità e scarsa razionalità, dal momento che determinano diverse discipline per dati non sempre facilmente distinguibili in base alle indicazioni normative, le quali impongono distinzioni sulla base di criteri non totalmente ragionevoli[99].

---

[96] Art. 3, n. 33 della proposta. Il considerando 7) chiarisce che tale nozione dovrebbe essere interpretata in modo coerente con la nozione di dati biometrici contenuta nel GDPR e nella LED.

[97] Alcuni chiarimenti sulla nozione si possono trovare nelle linee guida dell'EDPB sul trattamento dei dati personali attraverso dispositivi video, in cui si evidenziano tre componenti sulla base della definizione all'articolo 4 e per l'applicazione dell'art. 9 del GDPR: la «natura dei dati» (che devono essere relativi alle caratteristiche fisiche, fisiologiche o comportamentali di una persona fisica), i «mezzi e modalità del trattamento» (i dati devono essere ottenuti da un trattamento tecnico specifico), la «finalità del trattamento» (identificazione univoca della persona). Cfr. EDPB, Linee guida 3/2019 sul trattamento dei dati personali attraverso dispositivi video, 29 gennaio 2020, punto 76.

[98] Così chiarisce espressamente il considerando 51) del GDPR.

[99] Si veda ad esempio E. J. Kindt, *Having yes, using no? About the new legal regime for biometric data*, in *Computer Law and Security Review*, 3, 2018, 12 ss., in cui si identificano quattro discipline diverse all'interno del GDPR per dati quali immagini facciali; E. J. Kindt, *A First Attempt at Regulating Biometric Data in the European Union*, in A Kak (a cura di), *Regulating Biometrics*, *Global Approaches and Urgent Questions*, AI Now Institute, disponibile su *https://ainowinstitute.org/regulatingbiometrics.html*, 2020, 65 ss., in cui si dà atto anche del contrasto tra la protezione prevista per i dati biometrici nel GDPR (fondata sul trattamento dei dati ai fini di identificazione ai sensi dell'art. 9) e quella prevista dalla Corte europea dei diritti umani, che invece prescinde dal trattamento specifico dei dati. In T. Christakis, K. Bannelier, C. Castelluccia, D. Le Métayer, *Mapping the Use of Facial Recognition in Public Spaces in Europe – Part 1: A Quest for Clarity: Unpicking the 'Catch-All' Term"*, Report of the AI - Regulation Chair MIAI, 2022, disponibile su https://ai-regulation.com/facial-recognition-in-europe-part-1/, 19 ss. si dà atto anche della presentazione di emendamenti nel corso dell'esame del Parlamento europeo sulla proposta di regolamento in materia di IA, ai fini di superare la definizione di dati biometrici. Cfr. anche P. De Herta, G. Bouchagiar, *Visual and biometric surveillance in the EU. Saying 'no' to mass surveillance practices?*, in *Information Polity*, 2022, 2, 198 ss., Sulla ricostruzione della nozione cfr. anche C. Jasserand, *Legal Nature of Biometric Data: From 'Generic' Personal Data to Sensitive Data: Which Changes Does the New Data Protection Framework Introduce?*, in *European Data Protection Law Review*, 2016, 3, 297; E.J. Kindt, *Biometric data processing: Is the legislator keeping up or just keeping up appearances?*, in G. Gonzalez Fuster, R. Van Brakel, P. De Hert (a cura di), *Research Handbook on Privacy and Data Protection Law Values, Norms and Global Politics*, Cheltenham, 2022, 371 ss.; G. Gonzalez Fuster, M.A. Nadolna Peeters, *Person identification, human rights and ethical principles: Rethinking biometrics in the era of artificial intelligence*, European Parliament, 2021, 24 ss.

7. *Dati biometrici tra GDPR e regolamento sull'IA: due esempi concreti*

La connessione, illustrata nel paragrafo precedente, tra la proposta di regolamento sull'IA e la disciplina sulla protezione dei dati personali nel particolare ambito dei dati biometrici permette un esercizio: sperimentare l'applicazione delle categorie previste dalla proposta su tecnologie già esistenti e già oggetto di attenzione da parte delle autorità nazionali per la protezione dei dati personali.

Tra i tanti esempi possibili, ne prendiamo in considerazione due: una tecnologia di riconoscimento facciale e una di analisi della voce.

Come esempio di tecnologia di riconoscimento facciale, possiamo considerare il ben noto caso di Clearview AI, società statunitense che nell'ultimo biennio è stata sanzionata da diversi garanti degli Stati europei[100] per violazione del GDPR. La tecnologia di Clearview AI si basa sulla disponibilità di un (enorme) *database* di immagini di persone, creato dalla società attraverso lo *scraping* di immagini soprattutto da social network, associati a informazioni (metadati) sulle immagini ricavate con lo stesso metodo. Dalle immagini si ottengono rappresentazioni vettoriali del volto che riproducono le caratteristiche identificative delle persone. Il modello che si ricava viene indicizzato tramite *hashing* così da poter essere oggetto di ricerca[101]. Con questo metodo, il *database* di immagini può essere utilizzato come motore di ricerca ai fini dell'identificazione di persone: l'immagine della persona che si vuole cercare (*probe image* – immagine sonda[102]) viene a sua volta trasformata in un modello vettoriale e sottoposta a comparazione con i modelli indicizzati nel *database*, individuando le immagini corrispondenti (verifica *one-to-many*[103]).

---

[100] Tra gli altri, il garante francese (Commission Nationale de l'Informatique et des Libertés Décision MED-2021-134 del 26 novembre 2021), il garante italiano (Garante per la Protezione dei dati Personali, Ordinanza ingiunzione nei confronti di Clearview AI - 10 febbraio 2022 [9751362]), il garante del Regno Unito (UK Information Commissioner's Office, decisione del 26 maggio 2022), il garante greco (Αρχή Προστασίας Δεδομένων, decisione 13 luglio 2022).

[101] Cfr. Garante per la Protezione dei dati Personali, ordinanza 10 febbraio 2022, cit., par. 3.1.

[102] Cfr. UK Information Commissioner's Office, decisione del 26 maggio 2022, *Monetary penalty notice*, parr. 26 e ss.

[103] La verifica *one-to-many* presuppone la comparazione tra (il modello di) una singola immagine e un *data set*; si contrappone alla verifica *one-to-one*, che riguarda la comparazione tra due (modelli di) immagini, usata ai fini di verificazione/autenticazione (ad es. per lo sblocco di smartphone o verifica identità alle frontiere). Sul punto cfr. G. Gonzalez Fuster, M.A. Nadolna Peeters, cit., 8. Per esempi pratici di questi due tipi di riconoscimento (*one-to-many* e *one-to-one*) e le loro implicazioni giuridiche si vedano anche le recenti Linee guida dell'EDPB sul riconoscimento facciale per le attività di contrasto – EDPB, Guidelines 05/2022 on the use of facial recognition technology in the area of law enforcement, 12 maggio 2022). Sul funzionamento della tecnologia di Clearview si veda anche l'indagine delle autorità garanti canadesi (*Joint investigation of Clearview AI, Inc.* by the Office of the Privacy Commissioner of Canada, the Commission d'accès à l'information du Québec, the Information and Privacy Commissioner for British Columbia, and the Information Privacy Commissioner of Alberta, 2 febbraio 2021). Per una dimostrazione pratica si può vedere il *reportage* della CNN del 2017 disponibile su https://www.youtube.com/watch?v=q-1bR3P9RAw.

Clearview adotta un algoritmo di *machine learning* per la creazione e comparazione dei modelli e la tecnologia può rientrare dunque nell'ambito della definizione di sistema di intelligenza artificiale prevista dalla proposta di regolamento[104].

I dati utilizzati dalla tecnologia possono invece essere considerati come dati biometrici? Come visto sopra, le fotografie di persone fisiche non possono essere considerate *tout court* come dati biometrici ai sensi della definizione del GDPR (ripresa dalla proposta di regolamento), in quanto è necessario un «trattamento tecnico specifico». Nel caso di Clearview, quindi, le immagini di persone fisiche identificabili (e i metadati associati) contenute nel *database* "raschiato" dal web sono dati personali ma non biometrici, mentre i vettori tramite cui la tecnologia ottiene il modello delle immagini nel *database* (e dell'immagine sonda da sottoporre a comparazione) sono da considerare come biometrici[105].

La tecnologia di Clearview è dunque basata (almeno in parte) su dati biometrici e può dirsi finalizzata all'identificazione di persone fisiche. Nel caso in cui l'identificazione avvenga "a distanza" può rientrare nella nozione di «identificazione biometrica remota» prevista dalla proposta di regolamento ed essere quindi considerata, a seconda dei casi, sistema ad alto rischio (sulla base delle previsioni dell'allegato III della proposta sopra richiamati) oppure tecnologia vietata, se utilizzata «in tempo reale», in «spazi accessibili al pubblico» e per «attività di contrasto». La classificazione ai sensi del regolamento non pregiudica però l'applicazione delle regole in tema di protezione di dati personali, in base alle quali le autorità europee hanno riscontrato l'illegittimità della tecnologia nel contesto europeo.

Il secondo esempio – decisamente meno noto – riguarda una tecnologia di analisi vocale oggetto di un provvedimento dell'autorità garante ungherese[106]. La tecnologia, utilizzata da un istituto finanziario, analizzava le telefonate registrate del servizio clienti dell'impresa attraverso un sistema di *machine learning*. Il sistema individuava in particolare lo "stato emozionale" della conversazione sulla base della voce, delle pause e del numero delle persone intervenute, con lo scopo (nel caso di specie) di selezionare i clienti particolarmente insoddisfatti da richiamare, prevenire future lamentele, migliorare in generale i processi interni di gestione delle chiamate al *call center*[107].

---

[104] Va rilevato comunque che le tecnologie elencate nell'Allegato I dello schema di regolamento comprendono anche nozioni piuttosto ampie, quali «approcci statistici, stima bayesiana, metodi di ricerca e ottimizzazione» (lett. c) allegato I proposta di regolamento sull'IA).

[105] Tra i vari provvedimenti delle autorità garanti europee, l'Information Commissioner's Office del Regno Unito si sofferma con chiarezza sul punto: UK Information Commissioner's Office, decisione del 26 maggio 2022, Monetary penalty notice, parr. 38-40.

[106] Autorità nazionale per la protezione dei dati e la libertà di informazione, provvedimento 22 settembre 2021 nei confronti di Budapest Bank Zrt.

[107] Cfr. paragrafo I.1.(2) del provvedimento citato.

Anche in questo caso, la tecnologia può rientrare nell'ambito di applicazione della nozione di «intelligenza artificiale» prevista dalla proposta. I dati trattati però non vengono considerati dall'autorità garante come dati biometrici ai sensi del GDPR, in quanto lo strumento di analisi vocale non viene considerato quale «trattamento tecnico specifico» che consenta la «identificazione univoca» della persona fisica[108].

In questo caso, dunque, almeno seguendo l'interpretazione del provvedimento da cui è tratto l'esempio, la tecnologia non può dirsi basata su "dati biometrici" e non potrebbe quindi rientrare nella nozione dei sistemi specificamente presi in considerazione dalla proposta di regolamento (e in particolare quello relativo al «sistema di riconoscimento delle emozioni» richiamato nel paragrafo *supra*). Questo esito è probabilmente non auspicabile considerata la *ratio* delle definizioni delle tecnologie "biometriche" prese in considerazione, ma non precluderebbe, anche in questo caso, l'applicazione della disciplina sulla protezione dei dati personali (nella specie, il GDPR). Il GDPR è stato infatti applicato nel provvedimento dell'*authority* nazionale, che ha considerato illegittimo il trattamento effettuato pur non ritenendolo inerente a dati biometrici e ha sanzionato l'impresa coinvolta. In particolare, è stato rilevato come l'impresa non avesse provveduto ad informare adeguatamente gli interessati sui trattamenti effettuati (secondo un principio di trasparenza analogo a quello su cui è basata la disciplina dei «sistemi di riconoscimento delle emozioni» nella proposta di regolamento, come illustrato *supra*).

8. *Conclusioni*

La proposta di regolamento della Commissione sull'IA ha l'obiettivo molto ambizioso di regolare – per la prima volta in modo organico – tecnologie in continua e rapidissima evoluzione, con caratteristiche potenzialmente molto diverse fra loro, incentivando l'innovazione e promuovendo un clima di fiducia nel mercato tecnologico[109]. In questo senso, nella logica propria del ravvicinamento delle legislazioni per lo sviluppo del mercato interno, la regolazione comune dei sistemi di intelligenza artificiale mira a favorire la certezza del diritto europeo per gli operatori e gli investitori.

La complessità aumenta se si considera che il *draft* di regolamento si deve coordinare con altri atti normativi proposti dalla Commissione, nell'ambito della strategia europea in materia di dati e di mercati digitali, oltre che con il dirit-

[108] Cfr. paragrafo III.1. (45) del provvedimento citato.
[109] Comunicazione della Commissione europea COM(2021) 206, cit., par. 1.4.

to esistente[110]. Sotto quest'ultimo profilo, la proposta di regolamento sull'IA si deve misurare con il GDPR, riferimento principale in materia di trattamento dei dati personali: anche se l'oggetto di regolazione immediato della proposta sono i "sistemi di IA", e dunque le tecnologie e non i dati, è evidente che gli ambiti in cui le due discipline saranno chiamate ad integrarsi fra loro sono molteplici. I sistemi di IA – come ripetuto più volte – hanno come componente essenziale i dati e risultano particolarmente delicati quando coinvolgono dati personali, categoria peraltro con confini resi incerti dalle stesse modalità di funzionamento e possibilità delle tecnologie intelligenti.

Come visto, la proposta sull'IA e il GDPR condividono alcuni aspetti delle modalità di regolazione su cui si fondano, quali la regolazione basata sul rischio e la autoregolazione (*self regulation*) attraverso codici di condotta, la certificazione dei sistemi adottati secondo standard tecnici internazionali[111]. La valorizzazione dell'autonoma valutazione dell'operatore risponde ad una logica di responsabilizzazione e *accountability* dello stesso centrale nel sistema del GDPR. Tuttavia, nel contesto della proposta, non c'è la centralità propria del GDPR per i diritti degli interessati: l'autonomia degli operatori potrebbe tradursi in un'analisi superficiale dei rischi, soprattutto per le tecnologie non esattamente ed espressamente incluse nelle norme, a detrimento dei diritti fondamentali coinvolti.

Un secondo profilo problematico riguarda i punti in cui lo schema di regolamento sull'IA rinvia (o non rinvia) al GDPR. Nei casi in cui la proposta rinvia al GDPR, come per l'identificazione di tecnologie biometriche, si riprendono definizioni già di per sé complesse che rischiano di portare all'irragionevole esclusione di alcune tecnologie dall'applicazione della nuova disciplina. Nei casi in cui invece non si rinvia chiaramente alle regole del GDPR, questo potrebbe essere letto come una possibilità di derogare alla disciplina sul trattamento dei dati personali, come evidenziato dal parere congiunto di EDPB e GEPD. Anche questo aspetto, specie se sommato alla specificità e all'abbondanza di eccezioni con cui vengono identificate le tecnologie vietate (*in primis*, come visto, quelle di riconoscimento facciale), rischia di determinare eccessiva libertà di impiego di tecnologie particolarmente invasive per i diritti fondamentali.

[110] Oltre agli "atti legislativi" citati *supra* nella nota 92, si può menzionare la recentissima proposta della Commissione per una direttiva sulla responsabilità da intelligenza artificiale (Comunicazione della Commissione europea del 28.9.2022, COM(2022) 496).

[111] Sulle forme di regolazione impiegate nella proposta di regolamento sull'AI si veda L. Ammannati, *Diritti fondamentali e* rule of law *per una intelligenza artificiale*, in *Riv. Trim. Dir. dell'Economia,* 3 (supplemento), 2021, 170 ss. Sulle modalità di regolazione dell'intelligenza artificiale (e la regolazione mediante strumenti di intelligenza artificiale) cfr. anche L. Ammannati, F. Di Porto, *L'intelligenza artificiale per la fornitura di servizi, di applicazioni e la produzione di regole: Digital Services Act, Digital Markets Act e Artificial Intelligence Act*, in A. Pajno, F. Donati, A. Perrucci (a cura di), *Intelligenza artificiale e diritto: una rivoluzione?*, Vol. 1, Bologna, 2022, 479 ss. e M. Macchia, A. Mascolo, *Intelligenza artificiale e regolazione*, A. Pajno, F. Donati, A. Perrucci (a cura di), cit., Vol. 2, 97 ss.

Sarebbe dunque auspicabile che il legislatore europeo migliorasse alcuni aspetti della proposta di regolamento. Da un lato, sarebbe utile definire in modo più chiaro i rapporti tra la proposta sull'IA e gli altri atti dedicati ai dati, assicurandosi di coordinare al meglio le diverse discipline e riaffermando la centralità della protezione dei dati personali. In questa – non semplice – opera di coordinamento potrebbero anche essere previste delle modifiche alla normativa sulla protezione dei dati personali (GDPR e LED), volte ancora una volta ad eliminare elementi di incertezza e ambiguità, come quelli che caratterizzano la definizione di dati biometrici. Dall'altro lato, il legislatore potrebbe individuare con più decisione e meno eccezioni le tecnologie da vietare e quelle ad alto rischio, affinché il sistema di regolazione nel suo complesso si traduca in un'effettiva tutela dei diritti coinvolti, soprattutto nel caso delle tecnologie più invasive.

Nuove sfide della tecnologia e gestione dei rischi nella proposta di regolamento europeo sull'intelligenza artificiale: set di training, algoritmi e qualificazione dei dati. Profili critici

"Big data" e "intelligenza artificiale" sono due concetti chiave che contrassegnano l'evoluzione del mondo digitale. La proposta di regolamento europeo in materia di intelligenza artificiale mira a conciliare lo sviluppo sicuro e affidabile dell'IA con il rispetto dei valori e dei diritti dei cittadini dell'UE. L'articolo esamina il sistema di gestione dei rischi legati all'uso delle nuove tecnologie, soffermandosi sulle problematiche connesse alla definizione del set di addestramento della macchina e alla qualificazione dei dati che rientrano nell'ambito di applicazione del regolamento, in relazione al GDPR, con specifico riguardo ai dati biometrici.

New challenges related to technology and risk-management in the proposal for European regulation on artificial intelligence: training sets, algorithms and data qualification. Critical profiles

"Big data" and "artificial intelligence" are two key concepts that mark the evolution of the digital world. The draft of European artificial intelligence regulation aims to reconcile AI's safe and reliable development with respect for the values and rights of EU citizens. The article examines the risk management system related to the use of new technologies, focusing on the problems concerning the definition of the machine training set and the qualification of the data in light of the regulation proposal and the GDPR, with specific regard to the biometric data.